Exp Mail EV335610941US
USAN 09/780,669

**PCT**

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

# INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C07K 14/47, C12N 15/12, 5/10, C12Q 1/68, A61K 38/17

A2

(11) International Publication Number: **WO 98/17687**

(43) International Publication Date: 30 April 1998 (30.04.98)

(21) International Application Number: PCT/US97/19590

(22) International Filing Date: 24 October 1997 (24.10.97)

(30) Priority Data:

| | | |
|---|---|---|
| 08/740,274 | 25 October 1996 (25.10.96) | US |
| Not furnished | 24 October 1997 (24.10.97) | US |

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

(71) Applicant: GENETICS INSTITUTE, INC. [US/US]; 87 CambridgePark Drive, Cambridge, MA 02140 (US).

(72) Inventors: JACOBS, Kenneth; 151 Beaumont Avenue, Newton, MA 02160 (US). MCCOY, John, M.; 56 Howard Street, Reading, MA 01867 (US). LAVALLIE, Edward, R.; 90 Green Meadow Drive, Tewksbury, MA 01876 (US). RACIE, Lisa, A.; 124 School Street, Acton, MA 01720 (US). MERBERG, David; 2 Orchard Drive, Acton, MA 01720 (US). TREACY, Maurice; 93 Walcott Road, Chestnut Hill, MA 02167 (US). SPAULDING, Vikki; 11 Meadowbank Road, Billerica, MA 01821 (US). AGOSTINO, Michael, J.; 26 Wolcott Avenue, Andover, MA 01810 (US).

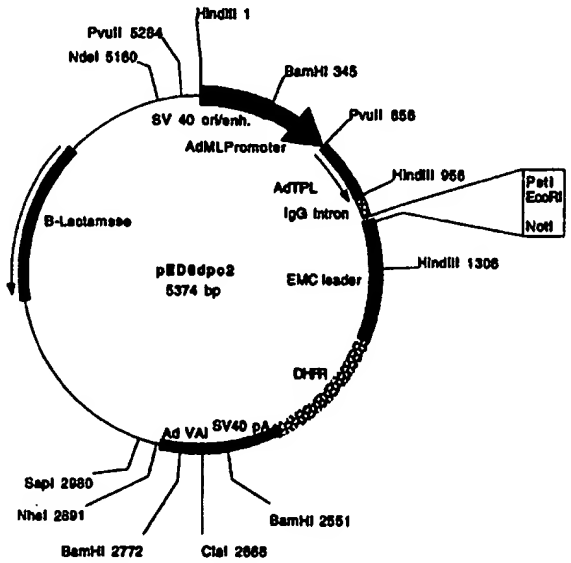
(74) Agent: SPRUNGER, Suzanne, A.; Genetics Institute, Inc., 87 CambridgePark Drive, Cambridge, MA 02140 (US).

**Published**
*Without international search report and to be republished upon receipt of that report.*

(54) Title: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM

(57) Abstract

Novel polynucleotides and the proteins encoded thereby are disclosed.




Plasmid name: pED6dpc2
Plasmid size: 5374 bp

Comments/References: pED6dpc2 is derived from pED6dpc1 by insertion of a new polylinker to facilitate cDNA cloning. SST cDNAs are cloned between EcoRI and NotI. pED vectors are described in Kaufman et al.(1991), NAR 19: 4485-4490.

*FOR THE PURPOSES OF INFORMATION ONLY*

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **AL** | Albania | **ES** | Spain | **LS** | Lesotho | **SI** | Slovenia |
| **AM** | Armenia | **FI** | Finland | **LT** | Lithuania | **SK** | Slovakia |
| **AT** | Austria | **FR** | France | **LU** | Luxembourg | **SN** | Senegal |
| **AU** | Australia | **GA** | Gabon | **LV** | Latvia | **SZ** | Swaziland |
| **AZ** | Azerbaijan | **GB** | United Kingdom | **MC** | Monaco | **TD** | Chad |
| **BA** | Bosnia and Herzegovina | **GE** | Georgia | **MD** | Republic of Moldova | **TG** | Togo |
| **BB** | Barbados | **GH** | Ghana | **MG** | Madagascar | **TJ** | Tajikistan |
| **BE** | Belgium | **GN** | Guinea | **MK** | The former Yugoslav Republic of Macedonia | **TM** | Turkmenistan |
| **BF** | Burkina Faso | **GR** | Greece | | | **TR** | Turkey |
| **BG** | Bulgaria | **HU** | Hungary | **ML** | Mali | **TT** | Trinidad and Tobago |
| **BJ** | Benin | **IE** | Ireland | **MN** | Mongolia | **UA** | Ukraine |
| **BR** | Brazil | **IL** | Israel | **MR** | Mauritania | **UG** | Uganda |
| **BY** | Belarus | **IS** | Iceland | **MW** | Malawi | **US** | United States of America |
| **CA** | Canada | **IT** | Italy | **MX** | Mexico | **UZ** | Uzbekistan |
| **CF** | Central African Republic | **JP** | Japan | **NE** | Niger | **VN** | Viet Nam |
| **CG** | Congo | **KE** | Kenya | **NL** | Netherlands | **YU** | Yugoslavia |
| **CH** | Switzerland | **KG** | Kyrgyzstan | **NO** | Norway | **ZW** | Zimbabwe |
| **CI** | Côte d'Ivoire | **KP** | Democratic People's Republic of Korea | **NZ** | New Zealand | | |
| **CM** | Cameroon | | | **PL** | Poland | | |
| **CN** | China | **KR** | Republic of Korea | **PT** | Portugal | | |
| **CU** | Cuba | **KZ** | Kazakstan | **RO** | Romania | | |
| **CZ** | Czech Republic | **LC** | Saint Lucia | **RU** | Russian Federation | | |
| **DE** | Germany | **LI** | Liechtenstein | **SD** | Sudan | | |
| **DK** | Denmark | **LK** | Sri Lanka | **SE** | Sweden | | |
| **EE** | Estonia | **LR** | Liberia | **SG** | Singapore | | |

# SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM

This application is a continuation-in-part of application Ser. No. 08/740,274, filed October 25, 1996.

## FIELD OF THE INVENTION

The present invention provides novel polynucleotides and proteins encoded by such polynucleotides, along with therapeutic, diagnostic and research utilities for these polynucleotides and proteins.

## BACKGROUND OF THE INVENTION

Technology aimed at the discovery of protein factors (including e.g., cytokines, such as lymphokines, interferons, CSFs and interleukins) has matured rapidly over the past decade. The now routine hybridization cloning and expression cloning techniques clone novel polynucleotides "directly" in the sense that they rely on information directly related to the discovered protein (i.e., partial DNA/amino acid sequence of the protein in the case of hybridization cloning; activity of the protein in the case of expression cloning). More recent "indirect" cloning techniques such as signal sequence cloning, which isolates DNA sequences based on the presence of a now well-recognized secretory leader sequence motif, as well as various PCR-based or low stringency hybridization cloning techniques, have advanced the state of the art by making available large numbers of DNA/amino acid sequences for proteins that are known to have biological activity by virtue of their secreted nature in the case of leader sequence cloning, or by virtue of the cell or tissue source in the case of PCR-based techniques. It is to these proteins and the polynucleotides encoding them that the present invention is directed.

## SUMMARY OF THE INVENTION

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 437 to nucleotide 1159;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 515 to nucleotide 1159;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 539 to nucleotide 1099;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone AR415_4 deposited under accession number ATCC 98232;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone AR415_4 deposited under accession number ATCC 98232;

(g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone AR415_4 deposited under accession number ATCC 98232;

(h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone AR415_4 deposited under accession number ATCC 98232;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:2;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:2 having biological activity;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:1 from nucleotide 437 to nucleotide 1159; the nucleotide sequence of SEQ ID NO:1 from nucleotide 515 to nucleotide 1159; the nucleotide sequence of SEQ ID NO:1 from

nucleotide 539 to nucleotide 1099; the nucleotide sequence of the full-length protein coding sequence of clone AR415_4 deposited under accession number ATCC 98232; or the nucleotide sequence of the mature protein coding sequence of clone AR415_4 deposited under accession number ATCC 98232. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone AR415_4 deposited under accession number ATCC 98232. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:2 from amino acid 51 to amino acid 221.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:1.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:2;

(b) the amino acid sequence of SEQ ID NO:2 from amino acid 51 to amino acid 221;

(c) fragments of the amino acid sequence of SEQ ID NO:2; and

(d) the amino acid sequence encoded by the cDNA insert of clone AR415_4 deposited under accession number ATCC 98232;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:2 or the amino acid sequence of SEQ ID NO:2 from amino acid 51 to amino acid 221.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 59 to nucleotide 376;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 179 to nucleotide 376;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone AS63_29 deposited under accession number ATCC 98232;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone AS63_29 deposited under accession number ATCC 98232;

(f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone AS63_29 deposited under accession number ATCC 98232;

(g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone AS63_29 deposited under accession number ATCC 98232;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:4;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:4 having biological activity;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:3 from nucleotide 59 to nucleotide 376; the nucleotide sequence of SEQ ID NO:3 from nucleotide 179 to nucleotide 376; the nucleotide sequence of the full-length protein coding sequence of clone AS63_29 deposited under accession number ATCC 98232; or the nucleotide sequence of the mature protein coding sequence of clone AS63_29 deposited under accession number ATCC 98232. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone AS63_29 deposited under accession number ATCC 98232. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:4 from amino acid 1 to amino acid 91.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:3 or SEQ ID NO:5.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:4;

(b) the amino acid sequence of SEQ ID NO:4 from amino acid 1 to amino acid 91;

(c) fragments of the amino acid sequence of SEQ ID NO:4; and

(d) the amino acid sequence encoded by the cDNA insert of clone AS63_29 deposited under accession number ATCC 98232;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:4 or the amino acid sequence of SEQ ID NO:4 from amino acid 1 to amino acid 91.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:6;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:6 from nucleotide 198 to nucleotide 2039;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:6 from nucleotide 490 to nucleotide 809;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone AY304_14 deposited under accession number ATCC xxxxx;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone AY304_14 deposited under accession number ATCC xxxxx;

(f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone AY304_14 deposited under accession number ATCC xxxxx;

(g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone AY304_14 deposited under accession number ATCC xxxxx;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:7;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:7 having biological activity;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:6 from nucleotide 198 to nucleotide 2039; the nucleotide sequence of SEQ ID NO:6 from nucleotide 490 to nucleotide 809; the nucleotide sequence of the full-length protein coding sequence of clone AY304_14 deposited under accession number ATCC xxxxx; or the nucleotide sequence of the mature protein coding sequence of clone AY304_14 deposited under accession number ATCC xxxxx. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone AY304_14 deposited under accession number ATCC xxxxx. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:7 from amino acid 126 to amino acid 204 or a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:7 from amino acid 106 to amino acid 204.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:6.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:7;

(b) the amino acid sequence of SEQ ID NO:7 from amino acid 126 to amino acid 204;

(c) the amino acid sequence of SEQ ID NO:7 from amino acid 106 to amino acid 204;

(d) fragments of the amino acid sequence of SEQ ID NO:7; and

(e) the amino acid sequence encoded by the cDNA insert of clone AY304_14 deposited under accession number ATCC xxxxx;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:7; the amino acid sequence of SEQ ID NO:7 from amino acid 126 to amino acid 204; or the amino acid sequence of SEQ ID NO:7 from amino acid 106 to amino acid 204.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:8;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:8 from nucleotide 102 to nucleotide 2027;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:8 from nucleotide 1902 to nucleotide 2027;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:8 from nucleotide 1 to nucleotide 431;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone BG160_1 deposited under accession number ATCC 98232;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone BG160_1 deposited under accession number ATCC 98232;

(g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone BG160_1 deposited under accession number ATCC 98232;

(h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone BG160_1 deposited under accession number ATCC 98232;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:9;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:9 having biological activity;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:8 from nucleotide 102 to nucleotide 2027; the nucleotide sequence of SEQ ID NO:8 from nucleotide 1902 to nucleotide 2027; the nucleotide sequence of SEQ ID NO:8 from nucleotide 1 to nucleotide 431; the nucleotide sequence of the full-length protein coding sequence of clone BG160_1 deposited under accession number ATCC 98232; or the nucleotide sequence of the mature protein coding sequence of clone BG160_1 deposited

under accession number ATCC 98232. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone BG160_1 deposited under accession number ATCC 98232. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:9 from amino acid 1 to amino acid 110.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:8.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:9;

(b) the amino acid sequence of SEQ ID NO:9 from amino acid 1 to amino acid 110;

(c) fragments of the amino acid sequence of SEQ ID NO:9; and

(d) the amino acid sequence encoded by the cDNA insert of clone BG160_1 deposited under accession number ATCC 98232;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:9 or the amino acid sequence of SEQ ID NO:9 from amino acid 1 to amino acid 110.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:11;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:11 from nucleotide 566 to nucleotide 631;

(c) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone BO432_4 deposited under accession number ATCC 98232;

(d) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone BO432_4 deposited under accession number ATCC 98232;

(e) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone BO432_4 deposited under accession number ATCC 98232;

(f) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone BO432_4 deposited under accession number ATCC 98232;

(g) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:12;

(h) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:12 having biological activity;

(i) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(f) above;

(j) a polynucleotide which encodes a species homologue of the protein of (g) or (h) above ; and

(k) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(h).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:11 from nucleotide 566 to nucleotide 631; the nucleotide sequence of the full-length protein coding sequence of clone BO432_4 deposited under accession number ATCC 98232; or the nucleotide sequence of the mature protein coding sequence of clone BO432_4 deposited under accession number ATCC 98232. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone BO432_4 deposited under accession number ATCC 98232.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:11, SEQ ID NO:10 or SEQ ID NO:13 .

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:12;

(b) fragments of the amino acid sequence of SEQ ID NO:12; and

(c) the amino acid sequence encoded by the cDNA insert of clone BO432_4 deposited under accession number ATCC 98232;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:12.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:14;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:14 from nucleotide 45 to nucleotide 428;

(c) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone BO538_2 deposited under accession number ATCC 98232;

(d) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone BO538_2 deposited under accession number ATCC 98232;

(e) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone BO538_2 deposited under accession number ATCC 98232;

(f) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone BO538_2 deposited under accession number ATCC 98232;

(g) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:15;

(h) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:15 having biological activity;

(i) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(f) above;

(j) a polynucleotide which encodes a species homologue of the protein of (g) or (h) above ; and

(k) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(h).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:14 from nucleotide 45 to nucleotide 428; the nucleotide sequence of the full-length protein coding sequence of clone BO538_2 deposited under accession number ATCC 98232; or the nucleotide sequence of the mature protein coding sequence of clone BO538_2 deposited under accession number ATCC 98232. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone BO538_2 deposited under accession number ATCC 98232. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:15 from amino acid 52 to amino acid 128.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:14 or SEQ ID NO:16.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:15;

(b) the amino acid sequence of SEQ ID NO:15 from amino acid 52 to amino acid 128;

(c) fragments of the amino acid sequence of SEQ ID NO:15; and

(d) the amino acid sequence encoded by the cDNA insert of clone BO538_2 deposited under accession number ATCC 98232;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:15 or the amino acid sequence of SEQ ID NO:15 from amino acid 52 to amino acid 128.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17 from nucleotide 144 to nucleotide 566;

(c) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone BR595_4 deposited under accession number ATCC 98232;

(d) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone BR595_4 deposited under accession number ATCC 98232;

(e) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone BR595_4 deposited under accession number ATCC 98232;

(f) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone BR595_4 deposited under accession number ATCC 98232;

(g) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:18;

(h) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:18 having biological activity;

(i) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(f) above;

(j) a polynucleotide which encodes a species homologue of the protein of (g) or (h) above ; and

(k) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(h).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:17 from nucleotide 144 to nucleotide 566; the nucleotide sequence of the full-length protein coding sequence of clone BR595_4 deposited under accession number ATCC 98232; or the nucleotide sequence of the mature protein coding sequence of clone BR595_4 deposited under accession number ATCC 98232. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone BR595_4 deposited under accession number ATCC 98232. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:18 from amino acid 39 to amino acid 141.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:17 or SEQ ID NO:19.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:18;

(b) the amino acid sequence of SEQ ID NO:18 from amino acid 39 to amino acid 141;

(c) fragments of the amino acid sequence of SEQ ID NO:18; and

(d) the amino acid sequence encoded by the cDNA insert of clone BR595_4 deposited under accession number ATCC 98232;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:18 or the amino acid sequence of SEQ ID NO:18 from amino acid 39 to amino acid 141.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:20;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:20 from nucleotide 232 to nucleotide 1041;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:20 from nucleotide 460 to nucleotide 1041;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:20 from nucleotide 590 to nucleotide 1163;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CI490_2 deposited under accession number ATCC 98232;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CI490_2 deposited under accession number ATCC 98232;

(g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CI490_2 deposited under accession number ATCC 98232;

(h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CI490_2 deposited under accession number ATCC 98232;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:21;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:21 having biological activity;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:20 from nucleotide 232 to nucleotide 1041; the nucleotide sequence of SEQ ID NO:20 from nucleotide 460 to nucleotide 1041; the nucleotide sequence of SEQ ID NO:20 from nucleotide 590 to nucleotide 1163; the nucleotide sequence of the full-length protein coding sequence of clone CI490_2 deposited under accession number ATCC 98232; or the nucleotide sequence of the mature protein coding sequence of clone CI490_2 deposited under accession number ATCC 98232. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone CI490_2 deposited under accession number ATCC 98232. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein

comprising the amino acid sequence of SEQ ID NO:21 from amino acid 133 to amino acid 270.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:20.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:21;

(b) the amino acid sequence of SEQ ID NO:21 from amino acid 133 to amino acid 270;

(c) fragments of the amino acid sequence of SEQ ID NO:21; and

(d) the amino acid sequence encoded by the cDNA insert of clone CI490_2 deposited under accession number ATCC 98232;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:21 or the amino acid sequence of SEQ ID NO:21 from amino acid 133 to amino acid 270.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:22;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:22 from nucleotide 268 to nucleotide 624;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:22 from nucleotide 325 to nucleotide 624;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CI522_1 deposited under accession number ATCC 98232;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CI522_1 deposited under accession number ATCC 98232;

(f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CI522_1 deposited under accession number ATCC 98232;

(g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CI522_1 deposited under accession number ATCC 98232;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:23;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:23 having biological activity;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:22 from nucleotide 268 to nucleotide 624; the nucleotide sequence of SEQ ID NO:22 from nucleotide 325 to nucleotide 624; the nucleotide sequence of the full-length protein coding sequence of clone CI522_1 deposited under accession number ATCC 98232; or the nucleotide sequence of the mature protein coding sequence of clone CI522_1 deposited under accession number ATCC 98232. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone CI522_1 deposited under accession number ATCC 98232.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:22 or SEQ ID NO:24.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:23;

(b) fragments of the amino acid sequence of SEQ ID NO:23; and

(c) the amino acid sequence encoded by the cDNA insert of clone CI522_1 deposited under accession number ATCC 98232;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:23.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:25;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:25 from nucleotide 288 to nucleotide 713;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:25 from nucleotide 686 to nucleotide 968;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CN238_1 deposited under accession number ATCC 98232;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CN238_1 deposited under accession number ATCC 98232;

(f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CN238_1 deposited under accession number ATCC 98232;

(g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CN238_1 deposited under accession number ATCC 98232;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:26;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:26 having biological activity;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:25 from nucleotide 288 to nucleotide 713; the nucleotide sequence of SEQ ID NO:25 from nucleotide 686 to nucleotide 968; the nucleotide sequence of the full-length protein coding sequence of clone CN238_1 deposited under accession number ATCC 98232; or the nucleotide sequence of the mature protein coding sequence of clone CN238_1 deposited under accession number ATCC 98232. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone CN238_1 deposited under accession number ATCC 98232.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:25.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:26;

(b) fragments of the amino acid sequence of SEQ ID NO:26; and

(c) the amino acid sequence encoded by the cDNA insert of clone CN238_1 deposited under accession number ATCC 98232;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:26.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:27;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:27 from nucleotide 87 to nucleotide 1874;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:27 from nucleotide 452 to nucleotide 830;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CO390_1 deposited under accession number ATCC 98232;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CO390_1 deposited under accession number ATCC 98232;

(f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CO390_1 deposited under accession number ATCC 98232;

(g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CO390_1 deposited under accession number ATCC 98232;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:28;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:28 having biological activity;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:27 from nucleotide 87 to nucleotide 1874; the nucleotide sequence of SEQ ID NO:27 from nucleotide 452 to nucleotide 830; the nucleotide sequence of the full-length protein coding sequence of clone CO390_1 deposited under accession number ATCC 98232; or the nucleotide sequence of the mature protein coding sequence of clone CO390_1 deposited under accession number ATCC 98232. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone CO390_1 deposited under accession number ATCC 98232. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:28 from amino acid 140 to amino acid 248.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:27.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:28;

(b) the amino acid sequence of SEQ ID NO:28 from amino acid 140 to amino acid 248;

(c) fragments of the amino acid sequence of SEQ ID NO:28; and

(d) the amino acid sequence encoded by the cDNA insert of clone CO390_1 deposited under accession number ATCC 98232;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:28 or the amino acid sequence of SEQ ID NO:28 from amino acid 140 to amino acid 248.

In certain preferred embodiments, the polynucleotide is operably linked to an expression control sequence. The invention also provides a host cell, including bacterial, yeast, insect and mammalian cells, transformed with such polynucleotide compositions.

Processes are also provided for producing a protein, which comprise:

(a) growing a culture of the host cell transformed with such polynucleotide compositions in a suitable culture medium; and

(b) purifying the protein from the culture.

The protein produced according to such methods is also provided by the present invention. Preferred embodiments include those in which the protein produced by such process is a mature form of the protein.

Protein compositions of the present invention may further comprise a pharmaceutically acceptable carrier. Compositions comprising an antibody which specifically reacts with such protein are also provided by the present invention.

Methods are also provided for preventing, treating or ameliorating a medical condition which comprises administering to a mammalian subject a therapeutically effective amount of a composition comprising a protein of the present invention and a pharmaceutically acceptable carrier.

## BRIEF DESCRIPTION OF DRAWINGS

Fig. 1 is a schematic representation of the pED6 and pNOTs vectors used for deposit of clones disclosed herein.

## DETAILED DESCRIPTION

### ISOLATED PROTEINS AND POLYNUCLEOTIDES

Nucleotide and amino acid sequences, as presently determined, are reported below for each clone and protein disclosed in the present application. The nucleotide sequence of each clone can readily be determined by sequencing of the deposited clone in accordance with known methods. The predicted amino acid sequence (both full-length and mature) can then be determined from such nucleotide sequence. The amino acid sequence of the protein encoded by a particular clone can also be determined by expression of the clone in a suitable host cell, collecting the protein and determining its sequence. For each disclosed protein applicants have identified what they have determined to be the reading frame best identifiable with sequence information available at the time of filing.

As used herein a "secreted" protein is one which, when expressed in a suitable host cell, is transported across or through a membrane, including transport as a result of signal sequences in its amino acid sequence. "Secreted" proteins include without limitation proteins secreted wholly (e.g., soluble proteins) or partially (e.g. , receptors) from the cell

in which they are expressed. "Secreted" proteins also include without limitation proteins which are transported across the membrane of the endoplasmic reticulum.

Clone "AR415_4"

A polynucleotide of the present invention has been identified as clone "AR415_4". AR415_4 was isolated from a human adult retina cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. AR415_4 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "AR415_4 protein").

The nucleotide sequence of AR415_4 as presently determined is reported in SEQ ID NO:1. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the AR415_4 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:2. Amino acids 14 to 26 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 27, or are a transmembrane domain.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone AR415_4 should be approximately 1500 bp.

The nucleotide sequence disclosed herein for AR415_4 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. AR415_4 demonstrated at least some homology with sequences identified as AA100799 (zm26d01.s1 Stratagene pancreas (#937208) Homo sapiens cDNA clone 526753 3'), AA100852 (zm26d01.r1 Stratagene pancreas (#937208) Homo sapiens cDNA clone 526753 5' similar to SW CO02_HUMAN P19075 TUMOR-ASSOCIATED ANTIGEN CO-029), AA146605 (zo35c09.r1 Stratagene colon (#937204) Homo sapiens cDNA clone 588880 5' similar to SW:CO02_HUMAN P19075 TUMOR-ASSOCIATED ANTIGEN CO-029), AA224847 (nc33c12.s1 NCI CGAP Pr2 Homo sapiens cDNA clone 4079 similar to SW:CO02_HUMAN P19075 TUMOR-ASSOCIATED ANTIGEN CO-029), AA225191 (nc21h08.s1 NCI CGAP Pr1 Homo sapiens cDNA clone 2968), AA593864 (nn19f08.s1 NCI_CGAP_Co12 Homo sapiens cDNA clone IMAGE:1084359), D26483 (Mouse mRNA for PE31/TALLA. 3/ ), M33680 (Human 26-kDa cell surface protein TAPA-1 mRNA, complete cds), T14726 (Human CD53 antigen cDNA), and T23814 (Human gene signature HUMGS05723). The predicted amino acid sequence disclosed

herein for AR415_4 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted AR415_4 protein demonstrated at least some identity with sequences identified as D29808 (TALLA-1 [Homo sapiens]), M35252 (tumor-associated antigen [Homo sapiens]), and R22360 (CO-029 tumour associated antigen protein). Based upon homology, AR415_4 proteins and each homologous protein or peptide may share at least some activity. The TopPredII computer program predicts a potential transmembrane domain within the AR415_4 protein sequence centered around amino acid 100 of SEQ ID NO:2.

Clone "AS63_29"

A polynucleotide of the present invention has been identified as clone "AS63_29". AS63_29 was isolated from a human fetal brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. AS63_29 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "AS63_29 protein").

The nucleotide sequence of the 5' portion of AS63_29 as presently determined is reported in SEQ ID NO:3. What applicants presently believe is the proper reading frame for the coding region is indicated in SEQ ID NO:4. The predicted amino acid sequence of the AS63_29 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:4. Amino acids 28 to 40 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 41, or are a transmembrane domain. Additional nucleotide sequence from the 3' portion of AS63_29, including the polyA tail, is reported in SEQ ID NO:5.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone AS63_29 should be approximately 1700 bp.

The nucleotide sequence disclosed herein for AS63_29 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. AS63_29 demonstrated at least some homology with sequences identified as L26877 (Mus musculus (B20c) heavy chain immunoglobulin variable region gene), T09146 (EST07039 Homo sapiens cDNA clone HIBBP68 5' end), T23466 (seq3050 Homo sapiens cDNA clone Hy18-Ch13-Charon40-cDNA-100 3'), and W55739 (ma35f05.r1 Life Tech mouse brain Mus musculus cDNA clone 312705 5'). The predicted amino acid

sequence disclosed herein for AS63_29 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted AS63_29 protein demonstrated at least some identity with sequences identified as R04032 (Full length T4 encoded by plasmid pBG381). Based upon homology, AS63_29 proteins and each homologous protein or peptide may share at least some activity. The TopPredII computer program predicts a potential transmembrane domain within the AS63_29 protein sequence, near the amino terminus.

Clone "AY304_14"

A polynucleotide of the present invention has been identified as clone "AY304_14". AY304_14 was isolated from a human adult retina cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. AY304_14 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "AY304_14 protein").

The nucleotide sequence of AY304_14 as presently determined is reported in SEQ ID NO:6. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the AY304_14 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:7.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone AY304_14 should be approximately 2200 bp.

The nucleotide sequence disclosed herein for AY304_14 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. AY304_14 demonstrated at least some homology with sequences identified as AA127688 (zk92f05.s1 Soares pregnant uterus NbHPU Homo sapiens cDNA clone 490305 3'), AA179609 (zp49g11.r1 Stratagene HeLa cell s3 937216 Homo sapiens cDNA clone 612836 5'), AA276253 (vc40f05.r1 Barstead MPLRB1 Mus musculus cDNA clone 777057 5'), H15545 (ym27d04.s1 Homo sapiens cDNA clone 49495 3' similar to contains PTR5 repetitive element), L08441 (Human autonomously replicating sequence (ARS) mRNA), N34949 (yy49h09.s1 Homo sapiens cDNA clone 276929 3'), R48594 (yj65d07.s1 Homo sapiens cDNA clone 153613 3'), T21160 (Human gene signature HUMGS02466), U43284 (Cloning vector phGFP-S65T, complete sequence, green fluorescent protein (gfp) gene, complete cds), and Z45151 (H. sapiens partial cDNA

sequence; clone c-2hh04). The predicted amino acid sequence disclosed herein for AY304_14 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted AY304_14 protein demonstrated at least some identity with sequences identified as D86984 (similar to yeast adenylate cyclase (S56776) [Homo sapiens]), J01415 (cytochrome oxidase subunit 3 [Homo sapiens]), V00662 (cytochrome oxidase III [Homo sapiens]), and X68948 (envelope glycoprotein [Spleen focus-forming virus]). Based upon homology, AY304_14 proteins and each homologous protein or peptide may share at least some activity. The TopPredII computer program predicts two potential transmembrane domains within the AY304_14 protein sequence, one centered around amino acid 81 and another around amino acid 120 of SEQ ID NO:7.

Clone "BG160_1"

A polynucleotide of the present invention has been identified as clone "BG160_1". BG160_1 was isolated from a human adult brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. BG160_1 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "BG160_1 protein").

The nucleotide sequence of BG160_1 as presently determined is reported in SEQ ID NO:8. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the BG160_1 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:9. Amino acids 588 to 600 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 601, or are a transmembrane domain.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone BG160_1 should be approximately 2300 bp.

The nucleotide sequence disclosed herein for BG160_1 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. BG160_1 demonstrated at least some homology with sequences identified as A60021 (tropomyosin-related protein, neuronal - rat ;contains element MER27 repetitive element), AA081525 (zn20e02.r1 Stratagene neuroepithelium NT2RAMI 937234 Homo sapiens cDNA clone 547994 5'), AA092565 (il5773.seq.F Fetal heart, Lambda ZAP Express Homo sapiens cDNA 5'), D56138 (Human fetal brain cDNA 5'-end

GEN-416H11), D61090 (Human fetal brain cDNA 5'-end GEN-155A07), D61184 (Human fetal brain cDNA 5'-end GEN-165A01), L10335 (Homo sapiens neuro-endocrine-specific protein C (NSP) mRNA, complete cds), N21304 (yx53f07.s1 Homo sapiens cDNA clone 265477 3' similar to SP:A60021 A60021 TROPOMYOSIN-RELATED PROTEIN, NEURONAL), and W95814 (ze07f11.r1 Soares fetal heart NbHH19W Homo sapiens cDNA clone 358317 5' similar to PIR:A60021). The predicted amino acid sequence disclosed herein for BG160_1 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted BG160_1 protein demonstrated at least some identity with sequences identified as L10334 (neuroendocrine-specific protein B [Homo sapiens]), L10335 (neuroendocrine-specific protein C [Homo sapiens]). Based upon homology, BG160_1 proteins and each homologous protein or peptide may share at least some activity. The TopPredII computer program predicts three potential transmembrane domains within the BG160_1 protein sequence, centered around amino acids 84, 484, and 595 of SEQ ID NO:9.

Clone "BO432_4"

A polynucleotide of the present invention has been identified as clone "BO432_4". BO432_4 was isolated from a human adult retina cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. BO432_4 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "BO432_4 protein").

The nucleotide sequence of the 5' portion of BO432_4 as presently determined is reported in SEQ ID NO:10. An additional internal nucleotide sequence from BO432_4 as presently determined is reported in SEQ ID NO:11. What applicants believe is the proper reading frame and the predicted amino acid sequence encoded by such internal sequence is reported in SEQ ID NO:12. Additional nucleotide sequence from the 3' portion of BO432_4, including the polyA tail, is reported in SEQ ID NO:13.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone BO432_4 should be approximately 1700 bp.

The nucleotide sequence disclosed herein for BO432_4 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. BO432_4 demonstrated at least some homology with sequences

identified as AA283626 (zt15e09.s1 Soares NbHTGBC Homo sapiens cDNA clone 713224 3'), AA406486 (zv12g02.r1 Soares NhHMPu S1 Homo sapiens cDNA clone 753458 5' similar to WP F35G2.2 CE05809 E.COLI YCAC LIKE), AA570446 (nk62c12.s1 NCI_CGAP_Sch1 Homo sapiens cDNA clone IMAGE:1018102), N55855 (J3389F Homo sapiens cDNA clone J3389 5'), Q10613 (Rianodin receptor gene), T62691 (yc70d10.r1 Homo sapiens cDNA clone 86035 5'), and W90766 (zh79h04.s1 Soares fetal liver spleen 1NFLS S1 Homo sapiens cDNA clone 418327 3'). The predicted amino acid sequence disclosed herein for BO432_4 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted BO432_4 protein demonstrated at least some identity with sequences identified as Z69637 (F35G2.2 [Caenorhabditis elegans]). Based upon homology, BO432_4 proteins and each homologous protein or peptide may share at least some activity. The TopPredII computer program predicts a potential transmembrane domain at the amino terminus of the BO432_4 protein sequence. The BO432_4 protein may also contain the bacterial lysR family signature, a motif found in bacterial transcriptional regulators and which is possibly indicative of a helix-turn-helix structure.

Clone "BO538_2"

A polynucleotide of the present invention has been identified as clone "BO538_2". BO538_2 was isolated from a human adult retina cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. BO538_2 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "BO538_2 protein").

The nucleotide sequence of the 5' portion of BO538_2 as presently determined is reported in SEQ ID NO:14. What applicants presently believe is the proper reading frame for the coding region is indicated in SEQ ID NO:15. The predicted amino acid sequence of the BO538_2 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:15. Additional nucleotide sequence from the 3' portion of BO538_2, including the polyA tail, is reported in SEQ ID NO:16.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone BO538_2 should be approximately 3000 bp.

The nucleotide sequence disclosed herein for BO538_2 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. BO538_2 demonstrated at least some homology with sequences identified as AA503100 (ne44h01.s1 NCI_CGAP_Co3 Homo sapiens cDNA clone 900241), R44035 (yg21g09.s1 Homo sapiens cDNA clone 33167 3'), T21630 (Human gene signature HUMGS03066), and W64854 (me06d12.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA clone 386711 5' similar to PIR S40989 S40989 hypothetical protein F55H2.6 - Caenorhabditis elegans). The predicted amino acid sequence disclosed herein for BO538_2 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted BO538_2 protein demonstrated at least some identity with sequences identified as M60525 (nerve growth factor inducible protein [Rattus norvegicus]), R28916 (Type III procollagen), and Z27080 (F55H2.6 [Caenorhabditis elegans]). Based upon homology, BO538_2 proteins and each homologous protein or peptide may share at least some activity. The TopPredII computer program predicts two potential transmembrane domains within the BO538_2 protein sequence.

Clone "BR595_4"

A polynucleotide of the present invention has been identified as clone "BR595_4". BR595_4 was isolated from a human fetal kidney cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. BR595_4 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "BR595_4 protein").

The nucleotide sequence of the 5' portion of BR595_4 as presently determined is reported in SEQ ID NO:17. What applicants presently believe is the proper reading frame for the coding region is indicated in SEQ ID NO:18. The predicted amino acid sequence of the BR595_4 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:18. Additional nucleotide sequence from the 3' portion of BR595_4, including the polyA tail, is reported in SEQ ID NO:19.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone BR595_4 should be approximately 3000 bp.

The nucleotide sequence disclosed herein for BR595_4 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. BR595_4 demonstrated at least some homology with sequences identified as AA443742 (zw95b02.s1 Soares total fetus Nb2HF8 9w Homo sapiens cDNA clone 784683 3'), AA600820 (np45b08.s1 NCI_CGAP_Br1.1 Homo sapiens cDNA clone IMAGE:1129239), T19410 (Human gene signature HUMGS00435), W87465 (zh67c04.s1 Soares fetal liver spleen 1NFLS S1 Homo sapiens cDNA clone 417126 3'), and Z33587 (H. sapiens partial cDNA sequence; clone HEA89P; single read). Based upon homology, BR595_4 proteins and each homologous protein or peptide may share at least some activity.

Clone "CI490_2"

A polynucleotide of the present invention has been identified as clone "CI490_2". CI490_2 was isolated from a human adult brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. CI490_2 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "CI490_2 protein").

The nucleotide sequence of CI490_2 as presently determined is reported in SEQ ID NO:20. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the CI490_2 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:21. Amino acids 64 to 76 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 77, or are a transmembrane domain.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone CI490_2 should be approximately 1200 bp.

The nucleotide sequence disclosed herein for CI490_2 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. CI490_2 demonstrated at least some homology with sequences identified as H30751 (yo79a04.r1 Homo sapiens cDNA clone 184110 5'), H49766 (yo24f01.r1 Homo sapiens cDNA clone 178873 5' similar to SP:S19586 N-METHYL-D-ASPARTATE RECEPTOR GLUTAMATE-BINDING CHAIN), H51158 (yo32d04.r1 Homo sapiens cDNA clone 179623 5'), R85211 (yo41d11.s1 Homo sapiens cDNA clone

180501 3' similar to SP S19586 N-METHYL-D-ASPARTATE RECEPTOR GLUTAMATE-BINDING CHAIN), S19586 (N-METHYL-D-ASPARTATE RECEPTOR GLUTAMATE-BINDING CHAIN), S61973 (NMDA receptor glutamate-binding subunit [rats, mRNA, 1742 nt]), T01031 (Human leucine zipper protein-kinase cDNA sequence), and W56893 (zc01g05.r1 Soares parathyroid tumor NbHPA Homo sapiens cDNA clone 321080 5' similar to PIR S19586 S19586 N-methyl-D-aspartate receptor glutamate-binding chain - rat). The predicted amino acid sequence disclosed herein for CI490_2 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted CI490_2 protein demonstrated at least some identity with sequences identified as S61973 (NMDA receptor glutamate-binding subunit [rats, Peptide, 516 aa] [Rattus sp.]) and U08020 (collagen pro-alpha-1 type I chain [Mus musculus]). Based upon homology, CI490_2 proteins and each homologous protein or peptide may share at least some activity. The TopPredII computer program predicts six potential transmembrane domains within the CI490_2 protein sequence, with the most amino-terminal transmembrane domain centered around amino acid 77 of SEQ ID NO:21.

Clone "CI522_1"

A polynucleotide of the present invention has been identified as clone "CI522_1". CI522_1 was isolated from a human adult brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. CI522_1 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "CI522_1 protein").

The nucleotide sequence of the 5' portion of CI522_1 as presently determined is reported in SEQ ID NO:22. What applicants presently believe is the proper reading frame for the coding region is indicated in SEQ ID NO:23. The predicted amino acid sequence of the CI522_1 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:23. Amino acids 7 to 19 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 20, or are a transmembrane domain. Additional nucleotide sequence from the 3' portion of CI522_1, including the polyA tail, is reported in SEQ ID NO:24.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone CI522_1 should be approximately 1400 bp.

The nucleotide sequence disclosed herein for CI522_1 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. CI522_1 demonstrated at least some homology with sequences identified as AA028557 (mi18g05.r1 Soares mouse p3NMF19.5 Mus musculus cDNA clone 463928 5'), H32238 (EST107136 Rattus sp. cDNA 5' end), T33525 (EST58140 Homo sapiens cDNA 5' end similar to None), U66468 (Human cell growth regulator CGR11 mRNA, complete cds), and X00525 (Mouse 28S ribosomal RNA). The predicted amino acid sequence disclosed herein for CI522_1 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted CI522_1 protein demonstrated at least some identity with sequences identified as U66468 (cell growth regulator CGR11 [Homo sapiens]). Based upon homology, CI522_1 proteins and each homologous protein or peptide may share at least some activity.

Clone "CN238_1"

A polynucleotide of the present invention has been identified as clone "CN238_1". CN238_1 was isolated from a human fetal brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. CN238_1 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "CN238_1 protein").

The nucleotide sequence of CN238_1 as presently determined is reported in SEQ ID NO:25. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the CN238_1 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:26.

The nucleotide sequence disclosed herein for CN238_1 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. CN238_1 demonstrated at least some homology with sequences identified as AA044097 (zk51b02.r1 Soares pregnant uterus NbHPU Homo sapiens cDNA clone 486315 5'), AA044287 (zk51b02.s1 Soares pregnant uterus NbHPU Homo sapiens cDNA clone 486315 3'), AA045440 (zk67c03.s1 Soares pregnant uterus NbHPU Homo sapiens cDNA clone 487876 3'), AA143007 (zl48f01.r1 Soares pregnant uterus NbHPU Homo sapiens cDNA clone 505177 5'), D51196 (Human fetal brain cDNA 3'-end

GEN-016G05), D60310 (Human fetal brain cDNA 3'-end GEN-098A09), N69344 (yz43e04.s1 Homo sapiens cDNA clone 285822 3' similar to gb:K00558 TUBULIN ALPHA-1 CHAIN (HUMAN)), W22250 (64B8 Human retina cDNA Tsp509I-cleaved sublibrary Homo), and X01703 (Human gene for alpha-tubulin (b alpha 1)). The predicted amino acid sequence disclosed herein for CN238_1 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted CN238_1 protein demonstrated at least some identity with sequences identified as K00557 (alpha-tubulin [Homo sapiens]) and U51583 (zinc finger homeodomain enhancer-binding protein-1 [Rattus norvegicus]). Based upon homology, CN238_1 proteins and each homologous protein or peptide may share at least some activity.

Clone "CO390_1"

A polynucleotide of the present invention has been identified as clone "CO390_1". CO390_1 was isolated from a human adult brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. CO390_1 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "CO390_1 protein").

The nucleotide sequence of CO390_1 as presently determined is reported in SEQ ID NO:27. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the CO390_1 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:28.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone CO390_1 should be approximately 2300 bp.

The nucleotide sequence disclosed herein for CO390_1 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. CO390_1 demonstrated at least some homology with sequences identified as H84353 (yv85a11.r1 Homo sapiens cDNA clone 249500 5'), L35532 (Pan troglodytes Alu repeat region), N80616 (Genomic clone encoding SAP(Phe)), R53922 (yi03h10.s1 Homo sapiens cDNA clone 138211 3' similar to contains Alu repetitive element;contains TAR1 repetitive element), X75335 (H.sapiens Alu insertion in COL3A1 gene), X95882 (R.norvegicus mRNA for ATP ligand gated ion channel), and Y09561 (H.sapiens mRNA for P2X7 receptor). The predicted amino acid sequence disclosed

herein for CO390_1 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted CO390_1 protein demonstrated at least some identity with sequences identified as U45448 (P2x1 receptor [Homo sapiens]), W04216 (Rat superior cervical ganglion p2x receptor), X83688 (ATP receptor [Homo sapiens]), X95882 (P2X7 gene product [Rattus norvegicus]), and Y09561 (ATP receptor [Homo sapiens]). Based upon homology, CO390_1 proteins and each homologous protein or peptide may share at least some activity. The TopPredII computer program predicts a potential transmembrane domain within the CO390_1 protein sequence, centered around amino acid 249 of SEQ ID NO:28. The nucleotide sequence of CO390_1 may contain an Alu repetitive element.

Deposit of Clones

Clones AR415_4, AS63_29, BG160_1, BO432_4, BO538_2, BR595_4, CI490_2, CI522_1, CN238_1, CO390_1, and AY304_1 (an additional isolate of clone AY304_14) were deposited on October 25, 1996 with the American Type Culture Collection as an original deposit under the Budapest Treaty and were given the accession number ATCC 98232, from which each clone comprising a particular polynucleotide is obtainable. Clone AY304_14 wasdeposited on October 23, 1997 with the American Type Culture Collection as an original deposit under the Budapest Treaty and was given the accession number ATCC xxxxx. All restrictions on the availability to the public of the deposited material will be irrevocably removed upon the granting of the patent, except for the requirements specified in 37 C.F.R. § 1.808(b).

Each clone has been transfected into separate bacterial cells (*E. coli*) in this composite deposit. Each clone can be removed from the vector in which it was deposited by performing an EcoRI/NotI digestion (5' site, EcoRI; 3' site, NotI) to produce the appropriate fragment for such clone. Each clone was deposited in either the pED6 or pNOTs vector depicted in Fig. 1. The pED6dpc2 vector ("pED6") was derived from pED6dpc1 by insertion of a new polylinker to facilitate cDNA cloning (Kaufman *et al.*, 1991, *Nucleic Acids Res.* **19**: 4485-4490); the pNOTs vector was derived from pMT2 (Kaufman *et al.*, 1989, *Mol. Cell. Biol.* 9: 946-958) by deletion of the DHFR sequences, insertion of a new polylinker, and insertion of the M13 origin of replication in the ClaI site. In some instances, the deposited clone can become "flipped" (i.e., in the reverse orientation) in the deposited isolate. In such instances, the cDNA insert can still be isolated by digestion with EcoRI and NotI. However, NotI will then produce the 5' site

and EcoRI will produce the 3' site for placement of the cDNA in proper orientation for expression in a suitable vector. The cDNA may also be expressed from the vectors in which they were deposited.

Bacterial cells containing a particular clone can be obtained from the composite deposit as follows:

An oligonucleotide probe or probes should be designed to the sequence that is known for that particular clone. This sequence can be derived from the sequences provided herein, or from a combination of those sequences. The sequence of the oligonucleotide probe that was used to isolate each full-length clone is identified below, and should be most reliable in isolating the clone of interest.

| Clone | Probe Sequence |
|---|---|
| AR415_4 | SEQ ID NO:29 |
| AS63_29 | SEQ ID NO:30 |
| AY304_14 | SEQ ID NO:31 |
| BG160_1 | SEQ ID NO:32 |
| BO432_4 | SEQ ID NO:33 |
| BO538_2 | SEQ ID NO:34 |
| BR595_4 | SEQ ID NO:35 |
| CI490_2 | SEQ ID NO:36 |
| CI522_1 | SEQ ID NO:37 |
| CN238_1 | SEQ ID NO:38 |
| CO390_1 | SEQ ID NO:39 |

In the sequences listed above which include an N at position 2, that position is occupied in preferred probes/primers by a biotinylated phosphoaramidite residue rather than a nucleotide (such as , for example, that produced by use of biotin phosphoramidite (1-dimethoxytrityloxy-2-(N-biotinyl-4-aminobutyl)-propyl-3-O-(2-cyanoethyl)-(N,N-diisopropyl)-phosphoramadite) (Glen Research, cat. no. 10-1953)).

The design of the oligonucleotide probe should preferably follow these parameters:

(a) It should be designed to an area of the sequence which has the fewest ambiguous bases ("N's"), if any;

(b) It should be designed to have a $T_m$ of approx. 80 ° C (assuming 2° for each A or T and 4 degrees for each G or C).

The oligonucleotide should preferably be labeled with g-$^{32}$P ATP (specific activity 6000 Ci/mmole) and T4 polynucleotide kinase using commonly employed techniques for labeling oligonucleotides. Other labeling techniques can also be used. Unincorporated label should preferably be removed by gel filtration chromatography or other established methods. The amount of radioactivity incorporated into the probe should be quantitated by measurement in a scintillation counter. Preferably, specific activity of the resulting probe should be approximately 4e+6 dpm/pmole.

The bacterial culture containing the pool of full-length clones should preferably be thawed and 100 µl of the stock used to inoculate a sterile culture flask containing 25 ml of sterile L-broth containing ampicillin at 100 µg/ml. The culture should preferably be grown to saturation at 37°C, and the saturated culture should preferably be diluted in fresh L-broth. Aliquots of these dilutions should preferably be plated to determine the dilution and volume which will yield approximately 5000 distinct and well-separated colonies on solid bacteriological media containing L-broth containing ampicillin at 100 µg/ml and agar at 1.5% in a 150 mm petri dish when grown overnight at 37°C. Other known methods of obtaining distinct, well-separated colonies can also be employed.

Standard colony hybridization procedures should then be used to transfer the colonies to nitrocellulose filters and lyse, denature and bake them.

The filter is then preferably incubated at 65°C for 1 hour with gentle agitation in 6X SSC (20X stock is 175.3 g NaCl/liter, 88.2 g Na citrate/liter, adjusted to pH 7.0 with NaOH) containing 0.5% SDS, 100 µg/ml of yeast RNA, and 10 mM EDTA (approximately 10 mL per 150 mm filter). Preferably, the probe is then added to the hybridization mix at a concentration greater than or equal to 1e+6 dpm/mL. The filter is then preferably incubated at 65°C with gentle agitation overnight. The filter is then preferably washed in 500 mL of 2X SSC/0.5% SDS at room temperature without agitation, preferably followed by 500 mL of 2X SSC/0.1% SDS at room temperature with gentle shaking for 15 minutes. A third wash with 0.1X SSC/0.5% SDS at 65°C for 30 minutes to 1 hour is optional. The filter is then preferably dried and subjected to autoradiography for sufficient time to visualize the positives on the X-ray film. Other known hybridization methods can also be employed.

The positive colonies are picked, grown in culture, and plasmid DNA isolated using standard procedures. The clones can then be verified by restriction analysis, hybridization analysis, or DNA sequencing.

Fragments of the proteins of the present invention which are capable of exhibiting biological activity are also encompassed by the present invention. Fragments of the protein may be in linear form or they may be cyclized using known methods, for example, as described in H.U. Saragovi, *et al.*, Bio/Technology 10, 773-778 (1992) and in R.S. McDowell, *et al.*, J. Amer. Chem. Soc. 114, 9245-9253 (1992), both of which are incorporated herein by reference. Such fragments may be fused to carrier molecules such as immunoglobulins for many purposes, including increasing the valency of protein binding sites. For example, fragments of the protein may be fused through "linker" sequences to the Fc portion of an immunoglobulin. For a bivalent form of the protein, such a fusion could be to the Fc portion of an IgG molecule. Other immunoglobulin isotypes may also be used to generate such fusions. For example, a protein - IgM fusion would generate a decavalent form of the protein of the invention.

The present invention also provides both full-length and mature forms of the disclosed proteins. The full-length form of the such proteins is identified in the sequence listing by translation of the nucleotide sequence of each disclosed clone. The mature form of such protein may be obtained by expression of the disclosed full-length polynucleotide (preferably those deposited with ATCC) in a suitable mammalian cell or other host cell. The sequence of the mature form of the protein may also be determinable from the amino acid sequence of the full-length form.

The present invention also provides genes corresponding to the cDNA sequences disclosed herein. "Corresponding genes" are the regions of the genome that are transcribed to produce the mRNAs from which the cDNA sequences are derived and any contiguous regions of the genome necessary for the regulated expression of such genes, including but not limited to coding sequences, 5' and 3' untranslated regions, alternatively spliced exons, introns, promoters, enhancers, and silencer or suppressor elements. The corresponding genes can be isolated in accordance with known methods using the sequence information disclosed herein. Such methods include the preparation of probes or primers from the disclosed sequence information for identification and/or amplification of genes in appropriate genomic libraries or other sources of genomic materials.

Where the protein of the present invention is membrane-bound (e.g., is a receptor), the present invention also provides for soluble forms of such protein. In such forms part or all of the intracellular and transmembrane domains of the protein are deleted such that the protein is fully secreted from the cell in which it is expressed. The intracellular and transmembrane domains of proteins of the invention can be identified in accordance with known techniques for determination of such domains from sequence information.

Proteins and protein fragments of the present invention include proteins with amino acid sequence lengths that are at least 25%(more preferably at least 50%, and most preferably at least 75%) of the length of a disclosed protein and have at least 60% sequence identity (more preferably, at least 75% identity; most preferably at least 90% or 95% identity) with that disclosed protein, where sequence identity is determined by comparing the amino acid sequences of the proteins when aligned so as to maximize overlap and identity while minimizing sequence gaps. Also included in the present invention are proteins and protein fragments that contain a segment preferably comprising 8 or more (more preferably 20 or more, most preferably 30 or more) contiguous amino acids that shares at least 75% sequence identity (more preferably, at least 85% identity; most preferably at least 95% identity) with any such segment of any of the disclosed proteins.

Species homologs of the disclosed polynucleotides and proteins are also provided by the present invention. Species homologs may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source from the desired species.

The invention also encompasses allelic variants of the disclosed polynucleotides or proteins; that is, naturally-occurring alternative forms of the isolated polynucleotide which also encode proteins which are identical, homologous or related to that encoded by the polynucleotides .

The invention also includes polynucleotides with sequences complementary to those of the polynucleotides disclosed herein.

The present invention also includes polynucleotides capable of hybridizing under reduced stringency conditions, more preferably stringent conditions, and most preferably highly stringent conditions, to polynucleotides described herein. Examples of stringency conditions are shown in the table below: highly stringent conditions are those that are at least as stringent as, for example, conditions A-F; stringent conditions are at least as stringent as, for example, conditions G-L; and reduced stringency conditions are at least as stringent as, for example, conditions M-R.

| Stringency Condition | Polynucleotide Hybrid | Hybrid Length (bp)‡ | Hybridization Temperature and Buffer† | Wash Temperature and Buffer† |
|---|---|---|---|---|
| A | DNA:DNA | ≥ 50 | 65°C; 1xSSC -or- 42°C; 1xSSC, 50% formamide | 65°C; 0.3xSSC |
| B | DNA:DNA | <50 | $T_B$*; 1xSSC | $T_B$*; 1xSSC |
| C | DNA:RNA | ≥ 50 | 67°C; 1xSSC -or- 45°C; 1xSSC, 50% formamide | 67°C; 0.3xSSC |
| D | DNA:RNA | <50 | $T_D$*; 1xSSC | $T_D$*; 1xSSC |
| E | RNA:RNA | ≥ 50 | 70°C; 1xSSC -or- 50°C; 1xSSC, 50% formamide | 70°C; 0.3xSSC |
| F | RNA:RNA | <50 | $T_F$*; 1xSSC | $T_F$*; 1xSSC |
| G | DNA:DNA | ≥ 50 | 65°C; 4xSSC -or- 42°C; 4xSSC, 50% formamide | 65°C; 1xSSC |
| H | DNA:DNA | <50 | $T_H$*; 4xSSC | $T_H$*; 4xSSC |
| I | DNA:RNA | ≥ 50 | 67°C; 4xSSC -or- 45°C; 4xSSC, 50% formamide | 67°C; 1xSSC |
| J | DNA:RNA | <50 | $T_J$*; 4xSSC | $T_J$*; 4xSSC |
| K | RNA:RNA | ≥ 50 | 70°C; 4xSSC -or- 50°C; 4xSSC, 50% formamide | 67°C; 1xSSC |
| L | RNA:RNA | <50 | $T_L$*; 2xSSC | $T_L$*; 2xSSC |
| M | DNA:DNA | ≥ 50 | 50°C; 4xSSC -or- 40°C; 6xSSC, 50% formamide | 50°C; 2xSSC |
| N | DNA:DNA | <50 | $T_N$*; 6xSSC | $T_N$*; 6xSSC |
| O | DNA:RNA | ≥ 50 | 55°C; 4xSSC -or- 42°C; 6xSSC, 50% formamide | 55°C; 2xSSC |
| P | DNA:RNA | <50 | $T_P$*; 6xSSC | $T_P$*; 6xSSC |
| Q | RNA:RNA | ≥ 50 | 60°C; 4xSSC -or- 45°C; 6xSSC, 50% formamide | 60°C; 2xSSC |
| R | RNA:RNA | <50 | $T_R$*; 4xSSC | $T_R$*; 4xSSC |

‡: The hybrid length is that anticipated for the hybridized region(s) of the hybridizing polynucleotides. When hybridizing a polynucleotide to a target polynucleotide of unknown sequence, the hybrid length is assumed to be that of the hybridizing polynucleotide. When polynucleotides of known sequence are hybridized, the hybrid length can be determined by aligning the sequences of the polynucleotides and identifying the region or regions of optimal sequence complementarity.

†: SSPE (1xSSPE is 0.15M NaCl, 10mM $NaH_2PO_4$, and 1.25mM EDTA, pH 7.4) can be substituted for SSC (1xSSC is 0.15M NaCl and 15mM sodium citrate) in the hybridization and wash buffers; washes are performed for 15 minutes after hybridization is complete.

*$T_B$ - $T_R$: The hybridization temperature for hybrids anticipated to be less than 50 base pairs in length should be 5-10°C less than the melting temperature ($T_m$) of the hybrid, where $T_m$ is determined according to the following equations. For hybrids less than 18 base pairs in length, $T_m$(°C) = 2(# of A + T bases) + 4(# of G + C bases). For hybrids between 18 and 49 base pairs in length, $T_m$(°C) = 81.5 + 16.6($\log_{10}$[$Na^+$]) + 0.41(%G+C) - (600/N), where N is the number of bases in the hybrid, and [$Na^+$] is the concentration of sodium ions in the hybridization buffer ([$Na^+$] for 1xSSC = 0.165 M).

Additional examples of stringency conditions for polynucleotide hybridization are provided in Sambrook, J., E.F. Fritsch, and T. Maniatis, 1989, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, chapters 9 and 11, and *Current Protocols in Molecular Biology*, 1995, F.M. Ausubel et al., eds., John Wiley & Sons, Inc., sections 2.10 and 6.3-6.4, incorporated herein by reference.

Preferably, each such hybridizing polynucleotide has a length that is at least 25%(more preferably at least 50%, and most preferably at least 75%) of the length of the polynucleotide of the present invention to which it hybridizes, and has at least 60% sequence identity (more preferably, at least 75% identity; most preferably at least 90% or 95% identity) with the polynucleotide of the present invention to which it hybridizes, where sequence identity is determined by comparing the sequences of the hybridizing polynucleotides when aligned so as to maximize overlap and identity while minimizing sequence gaps.

The isolated polynucleotide of the invention may be operably linked to an expression control sequence such as the pMT2 or pED expression vectors disclosed in Kaufman *et al.*, Nucleic Acids Res. 19, 4485-4490 (1991), in order to produce the protein recombinantly. Many suitable expression control sequences are known in the art. General methods of expressing recombinant proteins are also known and are exemplified in R. Kaufman, Methods in Enzymology 185, 537-566 (1990). As defined herein "operably linked" means that the isolated polynucleotide of the invention and an expression control sequence are situated within a vector or cell in such a way that the protein is expressed by a host cell which has been transformed (transfected) with the ligated polynucleotide/expression control sequence.

A number of types of cells may act as suitable host cells for expression of the protein. Mammalian host cells include, for example, monkey COS cells, Chinese Hamster Ovary (CHO) cells, human kidney 293 cells, human epidermal A431 cells, human Colo205 cells, 3T3 cells, CV-1 cells, other transformed primate cell lines, normal diploid cells, cell strains derived from in vitro culture of primary tissue, primary explants, HeLa cells, mouse L cells, BHK, HL-60, U937, HaK or Jurkat cells.

Alternatively, it may be possible to produce the protein in lower eukaryotes such as yeast or in prokaryotes such as bacteria. Potentially suitable yeast strains include *Saccharomyces cerevisiae, Schizosaccharomyces pombe, Kluyveromyces* strains, *Candida*, or any yeast strain capable of expressing heterologous proteins. Potentially suitable bacterial strains include *Escherichia coli, Bacillus subtilis, Salmonella typhimurium*, or any bacterial

strain capable of expressing heterologous proteins. If the protein is made in yeast or bacteria, it may be necessary to modify the protein produced therein, for example by phosphorylation or glycosylation of the appropriate sites, in order to obtain the functional protein. Such covalent attachments may be accomplished using known chemical or enzymatic methods.

The protein may also be produced by operably linking the isolated polynucleotide of the invention to suitable control sequences in one or more insect expression vectors, and employing an insect expression system. Materials and methods for baculovirus/insect cell expression systems are commercially available in kit form from, *e.g.*, Invitrogen, San Diego, California, U.S.A. (the MaxBac® kit), and such methods are well known in the art, as described in Summers and Smith, Texas Agricultural Experiment Station Bulletin No. 1555 (1987), incorporated herein by reference. As used herein, an insect cell capable of expressing a polynucleotide of the present invention is "transformed."

The protein of the invention may be prepared by culturing transformed host cells under culture conditions suitable to express the recombinant protein. The resulting expressed protein may then be purified from such culture (i.e., from culture medium or cell extracts) using known purification processes, such as gel filtration and ion exchange chromatography. The purification of the protein may also include an affinity column containing agents which will bind to the protein; one or more column steps over such affinity resins as concanavalin A-agarose, heparin-toyopearl® or Cibacrom blue 3GA Sepharose®; one or more steps involving hydrophobic interaction chromatography using such resins as phenyl ether, butyl ether, or propyl ether; or immunoaffinity chromatography.

Alternatively, the protein of the invention may also be expressed in a form which will facilitate purification. For example, it may be expressed as a fusion protein, such as those of maltose binding protein (MBP), glutathione-S-transferase (GST) or thioredoxin (TRX). Kits for expression and purification of such fusion proteins are commercially available from New England BioLab (Beverly, MA), Pharmacia (Piscataway, NJ) and InVitrogen, respectively. The protein can also be tagged with an epitope and subsequently purified by using a specific antibody directed to such epitope. One such epitope ("Flag") is commercially available from Kodak (New Haven, CT).

Finally, one or more reverse-phase high performance liquid chromatography (RP-HPLC) steps employing hydrophobic RP-HPLC media, e.g., silica gel having pendant

methyl or other aliphatic groups, can be employed to further purify the protein. Some or all of the foregoing purification steps, in various combinations, can also be employed to provide a substantially homogeneous isolated recombinant protein. The protein thus purified is substantially free of other mammalian proteins and is defined in accordance with the present invention as an "isolated protein."

The protein of the invention may also be expressed as a product of transgenic animals, e.g., as a component of the milk of transgenic cows, goats, pigs, or sheep which are characterized by somatic or germ cells containing a nucleotide sequence encoding the protein.

The protein may also be produced by known conventional chemical synthesis. Methods for constructing the proteins of the present invention by synthetic means are known to those skilled in the art. The synthetically-constructed protein sequences, by virtue of sharing primary, secondary or tertiary structural and/or conformational characteristics with proteins may possess biological properties in common therewith, including protein activity. Thus, they may be employed as biologically active or immunological substitutes for natural, purified proteins in screening of therapeutic compounds and in immunological processes for the development of antibodies.

The proteins provided herein also include proteins characterized by amino acid sequences similar to those of purified proteins but into which modification are naturally provided or deliberately engineered. For example, modifications in the peptide or DNA sequences can be made by those skilled in the art using known techniques. Modifications of interest in the protein sequences may include the alteration, substitution, replacement, insertion or deletion of a selected amino acid residue in the coding sequence. For example, one or more of the cysteine residues may be deleted or replaced with another amino acid to alter the conformation of the molecule. Techniques for such alteration, substitution, replacement, insertion or deletion are well known to those skilled in the art (see, e.g., U.S. Patent No. 4,518,584). Preferably, such alteration, substitution, replacement, insertion or deletion retains the desired activity of the protein.

Other fragments and derivatives of the sequences of proteins which would be expected to retain protein activity in whole or in part and may thus be useful for screening or other immunological methodologies may also be easily made by those skilled in the art given the disclosures herein. Such modifications are believed to be encompassed by the present invention.

## USES AND BIOLOGICAL ACTIVITY

The polynucleotides and proteins of the present invention are expected to exhibit one or more of the uses or biological activities (including those associated with assays cited herein) identified below. Uses or activities described for proteins of the present invention may be provided by administration or use of such proteins or by administration or use of polynucleotides encoding such proteins (such as, for example, in gene therapies or vectors suitable for introduction of DNA).

### Research Uses and Utilities

The polynucleotides provided by the present invention can be used by the research community for various purposes. The polynucleotides can be used to express recombinant protein for analysis, characterization or therapeutic use; as markers for tissues in which the corresponding protein is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in disease states); as molecular weight markers on Southern gels; as chromosome markers or tags (when labeled) to identify chromosomes or to map related gene positions; to compare with endogenous DNA sequences in patients to identify potential genetic disorders; as probes to hybridize and thus discover novel, related DNA sequences; as a source of information to derive PCR primers for genetic fingerprinting; as a probe to "subtract-out" known sequences in the process of discovering other novel polynucleotides; for selecting and making oligomers for attachment to a "gene chip" or other support, including for examination of expression patterns; to raise anti-protein antibodies using DNA immunization techniques; and as an antigen to raise anti-DNA antibodies or elicit another immune response. Where the polynucleotide encodes a protein which binds or potentially binds to another protein (such as, for example, in a receptor-ligand interaction), the polynucleotide can also be used in interaction trap assays (such as, for example, that described in Gyuris et al., Cell 75:791-803 (1993)) to identify polynucleotides encoding the other protein with which binding occurs or to identify inhibitors of the binding interaction.

The proteins provided by the present invention can similarly be used in assay to determine biological activity, including in a panel of multiple proteins for high-throughput screening; to raise antibodies or to elicit another immune response; as a reagent (including the labeled reagent) in assays designed to quantitatively determine levels of the protein (or its receptor) in biological fluids; as markers for tissues in which

the corresponding protein is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in a disease state); and, of course, to isolate correlative receptors or ligands. Where the protein binds or potentially binds to another protein (such as, for example, in a receptor-ligand interaction), the protein can be used to identify the other protein with which binding occurs or to identify inhibitors of the binding interaction. Proteins involved in these binding interactions can also be used to screen for peptide or small molecule inhibitors or agonists of the binding interaction.

Any or all of these research utilities are capable of being developed into reagent grade or kit format for commercialization as research products.

Methods for performing the uses listed above are well known to those skilled in the art. References disclosing such methods include without limitation "Molecular Cloning: A Laboratory Manual", 2d ed., Cold Spring Harbor Laboratory Press, Sambrook, J., E.F. Fritsch and T. Maniatis eds., 1989, and "Methods in Enzymology: Guide to Molecular Cloning Techniques", Academic Press, Berger, S.L. and A.R. Kimmel eds., 1987.

Nutritional Uses

Polynucleotides and proteins of the present invention can also be used as nutritional sources or supplements. Such uses include without limitation use as a protein or amino acid supplement, use as a carbon source, use as a nitrogen source and use as a source of carbohydrate. In such cases the protein or polynucleotide of the invention can be added to the feed of a particular organism or can be administered as a separate solid or liquid preparation, such as in the form of powder, pills, solutions, suspensions or capsules. In the case of microorganisms, the protein or polynucleotide of the invention can be added to the medium in or on which the microorganism is cultured.

Cytokine and Cell Proliferation/Differentiation Activity

A protein of the present invention may exhibit cytokine, cell proliferation (either inducing or inhibiting) or cell differentiation (either inducing or inhibiting) activity or may induce production of other cytokines in certain cell populations. Many protein factors discovered to date, including all known cytokines, have exhibited activity in one or more factor dependent cell proliferation assays, and hence the assays serve as a convenient confirmation of cytokine activity. The activity of a protein of the present invention is evidenced by any one of a number of routine factor dependent cell proliferation assays

for cell lines including, without limitation, 32D, DA2, DA1G, T10, B9, B9/11, BaF3, MC9/G, M+ (preB M+), 2E8, RB5, DA1, 123, T1165, HT2, CTLL2, TF-1, Mo7e and CMK.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for T-cell or thymocyte proliferation include without limitation those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Takai et al., J. Immunol. 137:3494-3500, 1986; Bertagnolli et al., J. Immunol. 145:1706-1712, 1990; Bertagnolli et al., Cellular Immunology 133:327-341, 1991; Bertagnolli, et al., J. Immunol. 149:3778-3783, 1992; Bowman et al., J. Immunol. 152: 1756-1761, 1994.

Assays for cytokine production and/or proliferation of spleen cells, lymph node cells or thymocytes include, without limitation, those described in: Polyclonal T cell stimulation, Kruisbeek, A.M. and Shevach, E.M. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 3.12.1-3.12.14, John Wiley and Sons, Toronto. 1994; and Measurement of mouse and human Interferon γ, Schreiber, R.D. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.8.1-6.8.8, John Wiley and Sons, Toronto. 1994.

Assays for proliferation and differentiation of hematopoietic and lymphopoietic cells include, without limitation, those described in: Measurement of Human and Murine Interleukin 2 and Interleukin 4, Bottomly, K., Davis, L.S. and Lipsky, P.E. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.3.1-6.3.12, John Wiley and Sons, Toronto. 1991; deVries et al., J. Exp. Med. 173:1205-1211, 1991; Moreau et al., Nature 336:690-692, 1988; Greenberger et al., Proc. Natl. Acad. Sci. U.S.A. 80:2931-2938, 1983; Measurement of mouse and human interleukin 6 - Nordan, R. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.6.1-6.6.5, John Wiley and Sons, Toronto. 1991; Smith et al., Proc. Natl. Acad. Sci. U.S.A. 83:1857-1861, 1986; Measurement of human Interleukin 11 - Bennett, F., Giannotti, J., Clark, S.C. and Turner, K. J. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.15.1 John Wiley and Sons, Toronto. 1991; Measurement of mouse and human Interleukin 9 - Ciarletta, A., Giannotti, J., Clark, S.C. and Turner, K.J. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.13.1, John Wiley and Sons, Toronto. 1991.

Assays for T-cell clone responses to antigens (which will identify, among others, proteins that affect APC-T cell interactions as well as direct T-cell effects by measuring proliferation and cytokine production) include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function; Chapter 6, Cytokines and their cellular receptors; Chapter 7, Immunologic studies in Humans); Weinberger et al., Proc. Natl. Acad. Sci. USA 77:6091-6095, 1980; Weinberger et al., Eur. J. Immun. 11:405-411, 1981; Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988.

Immune Stimulating or Suppressing Activity

A protein of the present invention may also exhibit immune stimulating or immune suppressing activity, including without limitation the activities for which assays are described herein. A protein may be useful in the treatment of various immune deficiencies and disorders (including severe combined immunodeficiency (SCID)), e.g., in regulating (up or down) growth and proliferation of T and/or B lymphocytes, as well as effecting the cytolytic activity of NK cells and other cell populations. These immune deficiencies may be genetic or be caused by viral (e.g., HIV) as well as bacterial or fungal infections, or may result from autoimmune disorders. More specifically, infectious diseases causes by viral, bacterial, fungal or other infection may be treatable using a protein of the present invention, including infections by HIV, hepatitis viruses, herpesviruses, mycobacteria, Leishmania spp., malaria spp. and various fungal infections such as candidiasis. Of course, in this regard, a protein of the present invention may also be useful where a boost to the immune system generally may be desirable, *i.e.*, in the treatment of cancer.

Autoimmune disorders which may be treated using a protein of the present invention include, for example, connective tissue disease, multiple sclerosis, systemic lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary inflammation, Guillain-Barre syndrome, autoimmune thyroiditis, insulin dependent diabetes mellitis, myasthenia gravis, graft-versus-host disease and autoimmune inflammatory eye disease. Such a protein of the present invention may also to be useful in the treatment of allergic reactions and conditions, such as asthma (particularly allergic asthma) or other respiratory problems. Other conditions, in which immune suppression is desired (including, for

example, organ transplantation), may also be treatable using a protein of the present invention.

Using the proteins of the invention it may also be possible to immune responses, in a number of ways. Down regulation may be in the form of inhibiting or blocking an immune response already in progress or may involve preventing the induction of an immune response. The functions of activated T cells may be inhibited by suppressing T cell responses or by inducing specific tolerance in T cells, or both. Immunosuppression of T cell responses is generally an active, non-antigen-specific, process which requires continuous exposure of the T cells to the suppressive agent. Tolerance, which involves inducing non-responsiveness or anergy in T cells, is distinguishable from immunosuppression in that it is generally antigen-specific and persists after exposure to the tolerizing agent has ceased. Operationally, tolerance can be demonstrated by the lack of a T cell response upon reexposure to specific antigen in the absence of the tolerizing agent.

Down regulating or preventing one or more antigen functions (including without limitation B lymphocyte antigen functions (such as , for example, B7)), *e.g.*, preventing high level lymphokine synthesis by activated T cells, will be useful in situations of tissue, skin and organ transplantation and in graft-versus-host disease (GVHD). For example, blockage of T cell function should result in reduced tissue destruction in tissue transplantation. Typically, in tissue transplants, rejection of the transplant is initiated through its recognition as foreign by T cells, followed by an immune reaction that destroys the transplant. The administration of a molecule which inhibits or blocks interaction of a B7 lymphocyte antigen with its natural ligand(s) on immune cells (such as a soluble, monomeric form of a peptide having B7-2 activity alone or in conjunction with a monomeric form of a peptide having an activity of another B lymphocyte antigen (*e.g.*, B7-1, B7-3) or blocking antibody), prior to transplantation can lead to the binding of the molecule to the natural ligand(s) on the immune cells without transmitting the corresponding costimulatory signal. Blocking B lymphocyte antigen function in this matter prevents cytokine synthesis by immune cells, such as T cells, and thus acts as an immunosuppressant. Moreover, the lack of costimulation may also be sufficient to anergize the T cells, thereby inducing tolerance in a subject. Induction of long-term tolerance by B lymphocyte antigen-blocking reagents may avoid the necessity of repeated administration of these blocking reagents. To achieve sufficient immunosuppression or

tolerance in a subject, it may also be necessary to block the function of a combination of B lymphocyte antigens.

The efficacy of particular blocking reagents in preventing organ transplant rejection or GVHD can be assessed using animal models that are predictive of efficacy in humans. Examples of appropriate systems which can be used include allogeneic cardiac grafts in rats and xenogeneic pancreatic islet cell grafts in mice, both of which have been used to examine the immunosuppressive effects of CTLA4Ig fusion proteins *in vivo* as described in Lenschow *et al.*, Science 257:789-792 (1992) and Turka *et al.*, Proc. Natl. Acad. Sci USA, *89*:11102-11105 (1992). In addition, murine models of GVHD (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 846-847) can be used to determine the effect of blocking B lymphocyte antigen function *in vivo* on the development of that disease.

Blocking antigen function may also be therapeutically useful for treating autoimmune diseases. Many autoimmune disorders are the result of inappropriate activation of T cells that are reactive against self tissue and which promote the production of cytokines and autoantibodies involved in the pathology of the diseases. Preventing the activation of autoreactive T cells may reduce or eliminate disease symptoms. Administration of reagents which block costimulation of T cells by disrupting receptor:ligand interactions of B lymphocyte antigens can be used to inhibit T cell activation and prevent production of autoantibodies or T cell-derived cytokines which may be involved in the disease process. Additionally, blocking reagents may induce antigen-specific tolerance of autoreactive T cells which could lead to long-term relief from the disease. The efficacy of blocking reagents in preventing or alleviating autoimmune disorders can be determined using a number of well-characterized animal models of human autoimmune diseases. Examples include murine experimental autoimmune encephalitis, systemic lupus erythmatosis in MRL/*lpr/lpr* mice or NZB hybrid mice, murine autoimmune collagen arthritis, diabetes mellitus in NOD mice and BB rats, and murine experimental myasthenia gravis (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 840-856).

Upregulation of an antigen function (preferably a B lymphocyte antigen function), as a means of up regulating immune responses, may also be useful in therapy. Upregulation of immune responses may be in the form of enhancing an existing immune response or eliciting an initial immune response. For example, enhancing an immune response through stimulating B lymphocyte antigen function may be useful in cases of

viral infection. In addition, systemic viral diseases such as influenza, the common cold, and encephalitis might be alleviated by the administration of stimulatory forms of B lymphocyte antigens systemically.

Alternatively, anti-viral immune responses may be enhanced in an infected patient by removing T cells from the patient, costimulating the T cells *in vitro* with viral antigen-pulsed APCs either expressing a peptide of the present invention or together with a stimulatory form of a soluble peptide of the present invention and reintroducing the *in vitro* activated T cells into the patient. Another method of enhancing anti-viral immune responses would be to isolate infected cells from a patient, transfect them with a nucleic acid encoding a protein of the present invention as described herein such that the cells express all or a portion of the protein on their surface, and reintroduce the transfected cells into the patient. The infected cells would now be capable of delivering a costimulatory signal to, and thereby activate, T cells *in vivo*.

In another application, up regulation or enhancement of antigen function (preferably B lymphocyte antigen function) may be useful in the induction of tumor immunity. Tumor cells (*e.g.*, sarcoma, melanoma, lymphoma, leukemia, neuroblastoma, carcinoma) transfected with a nucleic acid encoding at least one peptide of the present invention can be administered to a subject to overcome tumor-specific tolerance in the subject. If desired, the tumor cell can be transfected to express a combination of peptides. For example, tumor cells obtained from a patient can be transfected *ex vivo* with an expression vector directing the expression of a peptide having B7-2-like activity alone, or in conjunction with a peptide having B7-1-like activity and/or B7-3-like activity. The transfected tumor cells are returned to the patient to result in expression of the peptides on the surface of the transfected cell. Alternatively, gene therapy techniques can be used to target a tumor cell for transfection *in vivo*.

The presence of the peptide of the present invention having the activity of a B lymphocyte antigen(s) on the surface of the tumor cell provides the necessary costimulation signal to T cells to induce a T cell mediated immune response against the transfected tumor cells. In addition, tumor cells which lack MHC class I or MHC class II molecules, or which fail to reexpress sufficient amounts of MHC class I or MHC class II molecules, can be transfected with nucleic acid encoding all or a portion of (*e.g.*, a cytoplasmic-domain truncated portion) of an MHC class I $\alpha$ chain protein and $\beta_2$ microglobulin protein or an MHC class II $\alpha$ chain protein and an MHC class II $\beta$ chain protein to thereby express MHC class I or MHC class II proteins on the cell surface.

Expression of the appropriate class I or class II MHC in conjunction with a peptide having the activity of a B lymphocyte antigen (*e.g.*, B7-1, B7-2, B7-3) induces a T cell mediated immune response against the transfected tumor cell. Optionally, a gene encoding an antisense construct which blocks expression of an MHC class II associated protein, such as the invariant chain, can also be cotransfected with a DNA encoding a peptide having the activity of a B lymphocyte antigen to promote presentation of tumor associated antigens and induce tumor specific immunity. Thus, the induction of a T cell mediated immune response in a human subject may be sufficient to overcome tumor-specific tolerance in the subject.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for thymocyte or splenocyte cytotoxicity include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Herrmann et al., Proc. Natl. Acad. Sci. USA 78:2488-2492, 1981; Herrmann et al., J. Immunol. 128:1968-1974, 1982; Handa et al., J. Immunol. 135:1564-1572, 1985; Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Herrmann et al., Proc. Natl. Acad. Sci. USA 78:2488-2492, 1981; Herrmann et al., J. Immunol. 128:1968-1974, 1982; Handa et al., J. Immunol. 135:1564-1572, 1985; Takai et al., J. Immunol. 137:3494-3500, 1986; Bowmanet al., J. Virology 61:1992-1998; Takai et al., J. Immunol. 140:508-512, 1988; Bertagnolli et al., Cellular Immunology 133:327-341, 1991; Brown et al., J. Immunol. 153:3079-3092, 1994.

Assays for T-cell-dependent immunoglobulin responses and isotype switching (which will identify, among others, proteins that modulate T-cell dependent antibody responses and that affect Th1/Th2 profiles) include, without limitation, those described in: Maliszewski, J. Immunol. 144:3028-3033, 1990; and Assays for B cell function: *In vitro* antibody production, Mond, J.J. and Brunswick, M. In *Current Protocols in Immunology.* J.E.e.a. Coligan eds. Vol 1 pp. 3.8.1-3.8.16, John Wiley and Sons, Toronto. 1994.

Mixed lymphocyte reaction (MLR) assays (which will identify, among others, proteins that generate predominantly Th1 and CTL responses) include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter

7, Immunologic studies in Humans); Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Bertagnolli et al., J. Immunol. 149:3778-3783, 1992.

Dendritic cell-dependent assays (which will identify, among others, proteins expressed by dendritic cells that activate naive T-cells) include, without limitation, those described in: Guery et al., J. Immunol. 134:536-544, 1995; Inaba et al., Journal of Experimental Medicine 173:549-559, 1991; Macatonia et al., Journal of Immunology 154:5071-5079, 1995; Porgador et al., Journal of Experimental Medicine 182:255-260, 1995; Nair et al., Journal of Virology 67:4062-4069, 1993; Huang et al., Science 264:961-965, 1994; Macatonia et al., Journal of Experimental Medicine 169:1255-1264, 1989; Bhardwaj et al., Journal of Clinical Investigation 94:797-807, 1994; and Inaba et al., Journal of Experimental Medicine 172:631-640, 1990.

Assays for lymphocyte survival/apoptosis (which will identify, among others, proteins that prevent apoptosis after superantigen induction and proteins that regulate lymphocyte homeostasis) include, without limitation, those described in: Darzynkiewicz et al., Cytometry 13:795-808, 1992; Gorczyca et al., Leukemia 7:659-670, 1993; Gorczyca et al., Cancer Research 53:1945-1951, 1993; Itoh et al., Cell 66:233-243, 1991; Zacharchuk, Journal of Immunology 145:4037-4045, 1990; Zamai et al., Cytometry 14:891-897, 1993; Gorczyca et al., International Journal of Oncology 1:639-648, 1992.

Assays for proteins that influence early steps of T-cell commitment and development include, without limitation, those described in: Antica et al., Blood 84:111-117, 1994; Fine et al., Cellular Immunology 155:111-122, 1994; Galy et al., Blood 85:2770-2778, 1995; Toki et al., Proc. Nat. Acad Sci. USA 88:7548-7551, 1991.

Hematopoiesis Regulating Activity

A protein of the present invention may be useful in regulation of hematopoiesis and, consequently, in the treatment of myeloid or lymphoid cell deficiencies. Even marginal biological activity in support of colony forming cells or of factor-dependent cell lines indicates involvement in regulating hematopoiesis, e.g. in supporting the growth and proliferation of erythroid progenitor cells alone or in combination with other cytokines, thereby indicating utility, for example, in treating various anemias or for use in conjunction with irradiation/chemotherapy to stimulate the production of erythroid precursors and/or erythroid cells; in supporting the growth and proliferation of myeloid cells such as granulocytes and monocytes/macrophages (i.e., traditional CSF activity) useful, for example, in conjunction with chemotherapy to prevent or treat consequent

myelo-suppression; in supporting the growth and proliferation of megakaryocytes and consequently of platelets thereby allowing prevention or treatment of various platelet disorders such as thrombocytopenia, and generally for use in place of or complimentary to platelet transfusions; and/or in supporting the growth and proliferation of hematopoietic stem cells which are capable of maturing to any and all of the above-mentioned hematopoietic cells and therefore find therapeutic utility in various stem cell disorders (such as those usually treated with transplantation, including, without limitation, aplastic anemia and paroxysmal nocturnal hemoglobinuria), as well as in repopulating the stem cell compartment post irradiation/chemotherapy, either *in-vivo* or *ex-vivo* (i.e., in conjunction with bone marrow transplantation or with peripheral progenitor cell transplantation (homologous or heterologous)) as normal cells or genetically manipulated for gene therapy.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for proliferation and differentiation of various hematopoietic lines are cited above.

Assays for embryonic stem cell differentiation (which will identify, among others, proteins that influence embryonic differentiation hematopoiesis) include, without limitation, those described in: Johansson et al. Cellular Biology 15:141-151, 1995; Keller et al., Molecular and Cellular Biology 13:473-486, 1993; McClanahan et al., Blood 81:2903-2915, 1993.

Assays for stem cell survival and differentiation (which will identify, among others, proteins that regulate lympho-hematopoiesis) include, without limitation, those described in: Methylcellulose colony forming assays, Freshney, M.G. In *Culture of Hematopoietic Cells*. R.I. Freshney, *et al.* eds. Vol pp. 265-268, Wiley-Liss, Inc., New York, NY. 1994; Hirayama et al., Proc. Natl. Acad. Sci. USA 89:5907-5911, 1992; Primitive hematopoietic colony forming cells with high proliferative potential, McNiece, I.K. and Briddell, R.A. In *Culture of Hematopoietic Cells*. R.I. Freshney, *et al.* eds. Vol pp. 23-39, Wiley-Liss, Inc., New York, NY. 1994; Neben et al., Experimental Hematology 22:353-359, 1994; Cobblestone area forming cell assay, Ploemacher, R.E. In *Culture of Hematopoietic Cells*. R.I. Freshney, *et al.* eds. Vol pp. 1-21, Wiley-Liss, Inc.., New York, NY. 1994; Long term bone marrow cultures in the presence of stromal cells, Spooncer, E., Dexter, M. and Allen, T. In *Culture of Hematopoietic Cells*. R.I. Freshney, *et al.* eds. Vol pp. 163-179, Wiley-Liss, Inc., New York, NY. 1994; Long term culture initiating cell assay, Sutherland,

H.J. In *Culture of Hematopoietic Cells*. R.I. Freshney, *et al.* eds. Vol pp. 139-162, Wiley-Liss, Inc., New York, NY. 1994.

Tissue Growth Activity

A protein of the present invention also may have utility in compositions used for bone, cartilage, tendon, ligament and/or nerve tissue growth or regeneration, as well as for wound healing and tissue repair and replacement, and in the treatment of burns, incisions and ulcers.

A protein of the present invention, which induces cartilage and/or bone growth in circumstances where bone is not normally formed, has application in the healing of bone fractures and cartilage damage or defects in humans and other animals. Such a preparation employing a protein of the invention may have prophylactic use in closed as well as open fracture reduction and also in the improved fixation of artificial joints. *De novo* bone formation induced by an osteogenic agent contributes to the repair of congenital, trauma induced, or oncologic resection induced craniofacial defects, and also is useful in cosmetic plastic surgery.

A protein of this invention may also be used in the treatment of periodontal disease, and in other tooth repair processes. Such agents may provide an environment to attract bone-forming cells, stimulate growth of bone-forming cells or induce differentiation of progenitors of bone-forming cells. A protein of the invention may also be useful in the treatment of osteoporosis or osteoarthritis, such as through stimulation of bone and/or cartilage repair or by blocking inflammation or processes of tissue destruction (collagenase activity, osteoclast activity, etc.) mediated by inflammatory processes.

Another category of tissue regeneration activity that may be attributable to the protein of the present invention is tendon/ligament formation. A protein of the present invention, which induces tendon/ligament-like tissue or other tissue formation in circumstances where such tissue is not normally formed, has application in the healing of tendon or ligament tears, deformities and other tendon or ligament defects in humans and other animals. Such a preparation employing a tendon/ligament-like tissue inducing protein may have prophylactic use in preventing damage to tendon or ligament tissue, as well as use in the improved fixation of tendon or ligament to bone or other tissues, and in repairing defects to tendon or ligament tissue. De novo tendon/ligament-like tissue formation induced by a composition of the present invention contributes to the repair of

congenital, trauma induced, or other tendon or ligament defects of other origin, and is also useful in cosmetic plastic surgery for attachment or repair of tendons or ligaments. The compositions of the present invention may provide an environment to attract tendon- or ligament-forming cells, stimulate growth of tendon- or ligament-forming cells, induce differentiation of progenitors of tendon- or ligament-forming cells, or induce growth of tendon/ligament cells or progenitors *ex vivo* for return *in vivo* to effect tissue repair. The compositions of the invention may also be useful in the treatment of tendinitis, carpal tunnel syndrome and other tendon or ligament defects. The compositions may also include an appropriate matrix and/or sequestering agent as a carrier as is well known in the art.

The protein of the present invention may also be useful for proliferation of neural cells and for regeneration of nerve and brain tissue, *i.e.* for the treatment of central and peripheral nervous system diseases and neuropathies, as well as mechanical and traumatic disorders, which involve degeneration, death or trauma to neural cells or nerve tissue. More specifically, a protein may be used in the treatment of diseases of the peripheral nervous system, such as peripheral nerve injuries, peripheral neuropathy and localized neuropathies, and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome. Further conditions which may be treated in accordance with the present invention include mechanical and traumatic disorders, such as spinal cord disorders, head trauma and cerebrovascular diseases such as stroke. Peripheral neuropathies resulting from chemotherapy or other medical therapies may also be treatable using a protein of the invention.

Proteins of the invention may also be useful to promote better or faster closure of non-healing wounds, including without limitation pressure ulcers, ulcers associated with vascular insufficiency, surgical and traumatic wounds, and the like.

It is expected that a protein of the present invention may also exhibit activity for generation or regeneration of other tissues, such as organs (including, for example, pancreas, liver, intestine, kidney, skin, endothelium), muscle (smooth, skeletal or cardiac) and vascular (including vascular endothelium) tissue, or for promoting the growth of cells comprising such tissues. Part of the desired effects may be by inhibition or modulation of fibrotic scarring to allow normal tissue to regenerate. A protein of the invention may also exhibit angiogenic activity.

A protein of the present invention may also be useful for gut protection or regeneration and treatment of lung or liver fibrosis, reperfusion injury in various tissues, and conditions resulting from systemic cytokine damage.

A protein of the present invention may also be useful for promoting or inhibiting differentiation of tissues described above from precursor tissues or cells; or for inhibiting the growth of tissues described above.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for tissue generation activity include, without limitation, those described in: International Patent Publication No. WO95/16035 (bone, cartilage, tendon); International Patent Publication No. WO95/05846 (nerve, neuronal); International Patent Publication No. WO91/07491 (skin, endothelium ).

Assays for wound healing activity include, without limitation, those described in: Winter, Epidermal Wound Healing, pps. 71-112 (Maibach, HI and Rovee, DT, eds.), Year Book Medical Publishers, Inc., Chicago, as modified by Eaglstein and Mertz, J. Invest. Dermatol 71:382-84 (1978).

Activin/Inhibin Activity

A protein of the present invention may also exhibit activin- or inhibin-related activities. Inhibins are characterized by their ability to inhibit the release of follicle stimulating hormone (FSH), while activins and are characterized by their ability to stimulate the release of follicle stimulating hormone (FSH). Thus, a protein of the present invention, alone or in heterodimers with a member of the inhibin α family, may be useful as a contraceptive based on the ability of inhibins to decrease fertility in female mammals and decrease spermatogenesis in male mammals. Administration of sufficient amounts of other inhibins can induce infertility in these mammals. Alternatively, the protein of the invention, as a homodimer or as a heterodimer with other protein subunits of the inhibin-β group, may be useful as a fertility inducing therapeutic, based upon the ability of activin molecules in stimulating FSH release from cells of the anterior pituitary. See, for example, United States Patent 4,798,885. A protein of the invention may also be useful for advancement of the onset of fertility in sexually immature mammals, so as to increase the lifetime reproductive performance of domestic animals such as cows, sheep and pigs.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for activin/inhibin activity include, without limitation, those described in: Vale et al., Endocrinology 91:562-572, 1972; Ling et al., Nature 321:779-782, 1986; Vale et al., Nature 321:776-779, 1986; Mason et al., Nature 318:659-663, 1985; Forage et al., Proc. Natl. Acad. Sci. USA 83:3091-3095, 1986.

Chemotactic/Chemokinetic Activity

A protein of the present invention may have chemotactic or chemokinetic activity (e.g., act as a chemokine) for mammalian cells, including, for example, monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial cells. Chemotactic and chemokinetic proteins can be used to mobilize or attract a desired cell population to a desired site of action. Chemotactic or chemokinetic proteins provide particular advantages in treatment of wounds and other trauma to tissues, as well as in treatment of localized infections. For example, attraction of lymphocytes, monocytes or neutrophils to tumors or sites of infection may result in improved immune responses against the tumor or infecting agent.

A protein or peptide has chemotactic activity for a particular cell population if it can stimulate, directly or indirectly, the directed orientation or movement of such cell population. Preferably, the protein or peptide has the ability to directly stimulate directed movement of cells. Whether a particular protein has chemotactic activity for a population of cells can be readily determined by employing such protein or peptide in any known assay for cell chemotaxis.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for chemotactic activity (which will identify proteins that induce or prevent chemotaxis) consist of assays that measure the ability of a protein to induce the migration of cells across a membrane as well as the ability of a protein to induce the adhesion of one cell population to another cell population. Suitable assays for movement and adhesion include, without limitation, those described in: Current Protocols in Immunology, Ed by J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W.Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 6.12, Measurement of alpha and beta Chemokines 6.12.1-6.12.28; Taub et al. J. Clin. Invest. 95:1370-1376, 1995; Lind et al. APMIS 103:140-146, 1995; Muller et al Eur. J. Immunol. 25: 1744-1748; Gruber et al. J. of Immunol. 152:5860-5867, 1994; Johnston et al. J. of Immunol. 153: 1762-1768, 1994.

Hemostatic and Thrombolytic Activity

A protein of the invention may also exhibit hemostatic or thrombolytic activity. As a result, such a protein is expected to be useful in treatment of various coagulation disorders (including hereditary disorders, such as hemophilias) or to enhance coagulation and other hemostatic events in treating wounds resulting from trauma, surgery or other causes. A protein of the invention may also be useful for dissolving or inhibiting formation of thromboses and for treatment and prevention of conditions resulting therefrom (such as, for example, infarction of cardiac and central nervous system vessels (e.g., stroke).

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assay for hemostatic and thrombolytic activity include, without limitation, those described in: Linet et al., J. Clin. Pharmacol. 26:131-140, 1986; Burdick et al., Thrombosis Res. 45:413-419, 1987; Humphrey et al., Fibrinolysis 5:71-79 (1991); Schaub, Prostaglandins 35:467-474, 1988.

Receptor/Ligand Activity

A protein of the present invention may also demonstrate activity as receptors, receptor ligands or inhibitors or agonists of receptor/ligand interactions. Examples of such receptors and ligands include, without limitation, cytokine receptors and their ligands, receptor kinases and their ligands, receptor phosphatases and their ligands, receptors involved in cell-cell interactions and their ligands (including without limitation, cellular adhesion molecules (such as selectins, integrins and their ligands) and receptor/ligand pairs involved in antigen presentation, antigen recognition and development of cellular and humoral immune responses). Receptors and ligands are also useful for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interaction. A protein of the present invention (including, without limitation, fragments of receptors and ligands) may themselves be useful as inhibitors of receptor/ligand interactions.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for receptor-ligand activity include without limitation those described in:Current Protocols in Immunology, Ed by J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W.Strober, Pub. Greene Publishing Associates and

Wiley-Interscience (Chapter 7.28, Measurement of Cellular Adhesion under static conditions 7.28.1-7.28.22), Takai et al., Proc. Natl. Acad. Sci. USA 84:6864-6868, 1987; Bierer et al., J. Exp. Med. 168:1145-1156, 1988; Rosenstein et al., J. Exp. Med. 169:149-160 1989; Stoltenborg et al., J. Immunol. Methods 175:59-68, 1994; Stitt et al., Cell 80:661-670, 1995.

Anti-Inflammatory Activity

Proteins of the present invention may also exhibit anti-inflammatory activity. The anti-inflammatory activity may be achieved by providing a stimulus to cells involved in the inflammatory response, by inhibiting or promoting cell-cell interactions (such as, for example, cell adhesion), by inhibiting or promoting chemotaxis of cells involved in the inflammatory process, inhibiting or promoting cell extravasation, or by stimulating or suppressing production of other factors which more directly inhibit or promote an inflammatory response. Proteins exhibiting such activities can be used to treat inflammatory conditions including chronic or acute conditions), including without limitation inflammation associated with infection (such as septic shock, sepsis or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine or chemokine-induced lung injury, inflammatory bowel disease, Crohn's disease or resulting from over production of cytokines such as TNF or IL-1. Proteins of the invention may also be useful to treat anaphylaxis and hypersensitivity to an antigenic substance or material.

Cadherin/Tumor Invasion Suppressor Activity

Cadherins are calcium-dependent adhesion molecules that appear to play major roles during development, particularly in defining specific cell types. Loss or alteration of normal cadherin expression can lead to changes in cell adhesion properties linked to tumor growth and metastasis. Cadherin malfunction is also implicated in other human diseases, such as pemphigus vulgaris and pemphigus foliaceus (auto-immune blistering skin diseases), Crohn's disease, and some developmental abnormalities.

The cadherin superfamily includes well over forty members, each with a distinct pattern of expression. All members of the superfamily have in common conserved extracellular repeats (cadherin domains), but structural differences are found in other parts of the molecule. The cadherin domains bind calcium to form their tertiary structure and thus calcium is required to mediate their adhesion. Only a few amino acids in the

first cadherin domain provide the basis for homophilic adhesion; modification of this recognition site can change the specificity of a cadherin so that instead of recognizing only itself, the mutant molecule can now also bind to a different cadherin. In addition, some cadherins engage in heterophilic adhesion with other cadherins.

E-cadherin, one member of the cadherin superfamily, is expressed in epithelial cell types. Pathologically, if E-cadherin expression is lost in a tumor, the malignant cells become invasive and the cancer metastasizes. Transfection of cancer cell lines with polynucleotides expressing E-cadherin has reversed cancer-associated changes by returning altered cell shapes to normal, restoring cells' adhesiveness to each other and to their substrate, decreasing the cell growth rate, and drastically reducing anchorage-independent cell growth. Thus, reintroducing E-cadherin expression reverts carcinomas to a less advanced stage. It is likely that other cadherins have the same invasion suppressor role in carcinomas derived from other tissue types. Therefore, proteins of the present invention with cadherin activity, and polynucleotides of the present invention encoding such proteins, can be used to treat cancer. Introducing such proteins or polynucleotides into cancer cells can reduce or eliminate the cancerous changes observed in these cells by providing normal cadherin expression.

Cancer cells have also been shown to express cadherins of a different tissue type than their origin, thus allowing these cells to invade and metastasize in a different tissue in the body. Proteins of the present invention with cadherin activity, and polynucleotides of the present invention encoding such proteins, can be substituted in these cells for the inappropriately expressed cadherins, restoring normal cell adhesive properties and reducing or eliminating the tendency of the cells to metastasize.

Additionally, proteins of the present invention with cadherin activity, and polynucleotides of the present invention encoding such proteins, can used to generate antibodies recognizing and binding to cadherins. Such antibodies can be used to block the adhesion of inappropriately expressed tumor-cell cadherins, preventing the cells from forming a tumor elsewhere. Such an anti-cadherin antibody can also be used as a marker for the grade, pathological type, and prognosis of a cancer, i.e. the more progressed the cancer, the less cadherin expression there will be, and this decrease in cadherin expression can be detected by the use of a cadherin-binding antibody.

Fragments of proteins of the present invention with cadherin activity, preferably a polypeptide comprising a decapeptide of the cadherin recognition site, and polynucleotides of the present invention encoding such protein fragments, can also be used

to block cadherin function by binding to cadherins and preventing them from binding in ways that produce undesirable effects. Additionally, fragments of proteins of the present invention with cadherin activity, preferably truncated soluble cadherin fragments which have been found to be stable in the circulation of cancer patients, and polynucleotides encoding such protein fragments, can be used to disturb proper cell-cell adhesion.

Assays for cadherin adhesive and invasive suppressor activity include, without limitation, those described in: Hortsch et al. J Biol Chem 270 (32): 18809-18817, 1995; Miyaki et al. Oncogene 11: 2547-2552, 1995; Ozawa et al. Cell 63: 1033-1038, 1990.

Tumor Inhibition Activity

In addition to the activities described above for immunological treatment or prevention of tumors, a protein of the invention may exhibit other anti-tumor activities. A protein may inhibit tumor growth directly or indirectly (such as, for example, via ADCC). A protein may exhibit its tumor inhibitory activity by acting on tumor tissue or tumor precursor tissue, by inhibiting formation of tissues necessary to support tumor growth (such as, for example, by inhibiting angiogenesis), by causing production of other factors, agents or cell types which inhibit tumor growth, or by suppressing, eliminating or inhibiting factors, agents or cell types which promote tumor growth.

Other Activities

A protein of the invention may also exhibit one or more of the following additional activities or effects: inhibiting the growth, infection or function of, or killing, infectious agents, including, without limitation, bacteria, viruses, fungi and other parasites; effecting (suppressing or enhancing) bodily characteristics, including, without limitation, height, weight, hair color, eye color, skin, fat to lean ratio or other tissue pigmentation, or organ or body part size or shape (such as, for example, breast augmentation or diminution, change in bone form or shape); effecting biorhythms or caricadic cycles or rhythms; effecting the fertility of male or female subjects; effecting the metabolism, catabolism, anabolism, processing, utilization, storage or elimination of dietary fat, lipid, protein, carbohydrate, vitamins, minerals, cofactors or other nutritional factors or component(s); effecting behavioral characteristics, including, without limitation, appetite, libido, stress, cognition (including cognitive disorders), depression (including depressive disorders) and violent behaviors; providing analgesic effects or other pain reducing effects; promoting differentiation and growth of embryonic stem cells in lineages other than hematopoietic

lineages; hormonal or endocrine activity; in the case of enzymes, correcting deficiencies of the enzyme and treating deficiency-related diseases; treatment of hyperproliferative disorders (such as, for example, psoriasis); immunoglobulin-like activity (such as, for example, the ability to bind antigens or complement); and the ability to act as an antigen in a vaccine composition to raise an immune response against such protein or another material or entity which is cross-reactive with such protein.

## ADMINISTRATION AND DOSING

A protein of the present invention (from whatever source derived, including without limitation from recombinant and non-recombinant sources) may be used in a pharmaceutical composition when combined with a pharmaceutically acceptable carrier. Such a composition may also contain (in addition to protein and a carrier) diluents, fillers, salts, buffers, stabilizers, solubilizers, and other materials well known in the art. The term "pharmaceutically acceptable" means a non-toxic material that does not interfere with the effectiveness of the biological activity of the active ingredient(s). The characteristics of the carrier will depend on the route of administration. The pharmaceutical composition of the invention may also contain cytokines, lymphokines, or other hematopoietic factors such as M-CSF, GM-CSF, TNF, IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, IL-13, IL-14, IL-15, IFN, TNF0, TNF1, TNF2, G-CSF, Meg-CSF, thrombopoietin, stem cell factor, and erythropoietin. The pharmaceutical composition may further contain other agents which either enhance the activity of the protein or compliment its activity or use in treatment. Such additional factors and/or agents may be included in the pharmaceutical composition to produce a synergistic effect with protein of the invention, or to minimize side effects. Conversely, protein of the present invention may be included in formulations of the particular cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent to minimize side effects of the cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent.

A protein of the present invention may be active in multimers (e.g., heterodimers or homodimers) or complexes with itself or other proteins. As a result, pharmaceutical compositions of the invention may comprise a protein of the invention in such multimeric or complexed form.

The pharmaceutical composition of the invention may be in the form of a complex of the protein(s) of present invention along with protein or peptide antigens. The protein and/or peptide antigen will deliver a stimulatory signal to both B and T lymphocytes. B lymphocytes will respond to antigen through their surface immunoglobulin receptor. T lymphocytes will respond to antigen through the T cell receptor (TCR) following presentation of the antigen by MHC proteins. MHC and structurally related proteins including those encoded by class I and class II MHC genes on host cells will serve to present the peptide antigen(s) to T lymphocytes. The antigen components could also be supplied as purified MHC-peptide complexes alone or with co-stimulatory molecules that can directly signal T cells. Alternatively antibodies able to bind surface immunolgobulin and other molecules on B cells as well as antibodies able to bind the TCR and other molecules on T cells can be combined with the pharmaceutical composition of the invention.

The pharmaceutical composition of the invention may be in the form of a liposome in which protein of the present invention is combined, in addition to other pharmaceutically acceptable carriers, with amphipathic agents such as lipids which exist in aggregated form as micelles, insoluble monolayers, liquid crystals, or lamellar layers in aqueous solution. Suitable lipids for liposomal formulation include, without limitation, monoglycerides, diglycerides, sulfatides, lysolecithin, phospholipids, saponin, bile acids, and the like. Preparation of such liposomal formulations is within the level of skill in the art, as disclosed, for example, in U.S. Patent No. 4,235,871; U.S. Patent No. 4,501,728; U.S. Patent No. 4,837,028; and U.S. Patent No. 4,737,323, all of which are incorporated herein by reference.

As used herein, the term "therapeutically effective amount" means the total amount of each active component of the pharmaceutical composition or method that is sufficient to show a meaningful patient benefit, i.e., treatment, healing, prevention or amelioration of the relevant medical condition, or an increase in rate of treatment, healing, prevention or amelioration of such conditions. When applied to an individual active ingredient, administered alone, the term refers to that ingredient alone. When applied to a combination, the term refers to combined amounts of the active ingredients that result in the therapeutic effect, whether administered in combination, serially or simultaneously.

In practicing the method of treatment or use of the present invention, a therapeutically effective amount of protein of the present invention is administered to a mammal having a condition to be treated. Protein of the present invention may be

administered in accordance with the method of the invention either alone or in combination with other therapies such as treatments employing cytokines, lymphokines or other hematopoietic factors. When co-administered with one or more cytokines, lymphokines or other hematopoietic factors, protein of the present invention may be administered either simultaneously with the cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors, or sequentially. If administered sequentially, the attending physician will decide on the appropriate sequence of administering protein of the present invention in combination with cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors.

Administration of protein of the present invention used in the pharmaceutical composition or to practice the method of the present invention can be carried out in a variety of conventional ways, such as oral ingestion, inhalation, topical application or cutaneous, subcutaneous, intraperitoneal, parenteral or intravenous injection. Intravenous administration to the patient is preferred.

When a therapeutically effective amount of protein of the present invention is administered orally, protein of the present invention will be in the form of a tablet, capsule, powder, solution or elixir. When administered in tablet form, the pharmaceutical composition of the invention may additionally contain a solid carrier such as a gelatin or an adjuvant. The tablet, capsule, and powder contain from about 5 to 95% protein of the present invention, and preferably from about 25 to 90% protein of the present invention. When administered in liquid form, a liquid carrier such as water, petroleum, oils of animal or plant origin such as peanut oil, mineral oil, soybean oil, or sesame oil, or synthetic oils may be added. The liquid form of the pharmaceutical composition may further contain physiological saline solution, dextrose or other saccharide solution, or glycols such as ethylene glycol, propylene glycol or polyethylene glycol. When administered in liquid form, the pharmaceutical composition contains from about 0.5 to 90% by weight of protein of the present invention, and preferably from about 1 to 50% protein of the present invention.

When a therapeutically effective amount of protein of the present invention is administered by intravenous, cutaneous or subcutaneous injection, protein of the present invention will be in the form of a pyrogen-free, parenterally acceptable aqueous solution. The preparation of such parenterally acceptable protein solutions, having due regard to pH, isotonicity, stability, and the like, is within the skill in the art. A preferred

pharmaceutical composition for intravenous, cutaneous, or subcutaneous injection should contain, in addition to protein of the present invention, an isotonic vehicle such as Sodium Chloride Injection, Ringer's Injection, Dextrose Injection, Dextrose and Sodium Chloride Injection, Lactated Ringer's Injection, or other vehicle as known in the art. The pharmaceutical composition of the present invention may also contain stabilizers, preservatives, buffers, antioxidants, or other additives known to those of skill in the art.

The amount of protein of the present invention in the pharmaceutical composition of the present invention will depend upon the nature and severity of the condition being treated, and on the nature of prior treatments which the patient has undergone. Ultimately, the attending physician will decide the amount of protein of the present invention with which to treat each individual patient. Initially, the attending physician will administer low doses of protein of the present invention and observe the patient's response. Larger doses of protein of the present invention may be administered until the optimal therapeutic effect is obtained for the patient, and at that point the dosage is not increased further. It is contemplated that the various pharmaceutical compositions used to practice the method of the present invention should contain about 0.01 µg to about 100 mg (preferably about 0.1ng to about 10 mg, more preferably about 0.1 µg to about 1 mg) of protein of the present invention per kg body weight.

The duration of intravenous therapy using the pharmaceutical composition of the present invention will vary, depending on the severity of the disease being treated and the condition and potential idiosyncratic response of each individual patient. It is contemplated that the duration of each application of the protein of the present invention will be in the range of 12 to 24 hours of continuous intravenous administration. Ultimately the attending physician will decide on the appropriate duration of intravenous therapy using the pharmaceutical composition of the present invention.

Protein of the invention may also be used to immunize animals to obtain polyclonal and monoclonal antibodies which specifically react with the protein. Such antibodies may be obtained using either the entire protein or fragments thereof as an immunogen. The peptide immunogens additionally may contain a cysteine residue at the carboxyl terminus, and are conjugated to a hapten such as keyhole limpet hemocyanin (KLH). Methods for synthesizing such peptides are known in the art, for example, as in R.P. Merrifield, J. Amer.Chem.Soc. 85, 2149-2154 (1963); J.L. Krstenansky, *et al.*, FEBS Lett. 211, 10 (1987). Monoclonal antibodies binding to the protein of the invention may be useful diagnostic agents for the immunodetection of the protein. Neutralizing monoclonal

antibodies binding to the protein may also be useful therapeutics for both conditions associated with the protein and also in the treatment of some forms of cancer where abnormal expression of the protein is involved. In the case of cancerous cells or leukemic cells, neutralizing monoclonal antibodies against the protein may be useful in detecting and preventing the metastatic spread of the cancerous cells, which may be mediated by the protein.

For compositions of the present invention which are useful for bone, cartilage, tendon or ligament regeneration, the therapeutic method includes administering the composition topically, systematically, or locally as an implant or device. When administered, the therapeutic composition for use in this invention is, of course, in a pyrogen-free, physiologically acceptable form. Further, the composition may desirably be encapsulated or injected in a viscous form for delivery to the site of bone, cartilage or tissue damage. Topical administration may be suitable for wound healing and tissue repair. Therapeutically useful agents other than a protein of the invention which may also optionally be included in the composition as described above, may alternatively or additionally, be administered simultaneously or sequentially with the composition in the methods of the invention. Preferably for bone and/or cartilage formation, the composition would include a matrix capable of delivering the protein-containing composition to the site of bone and/or cartilage damage, providing a structure for the developing bone and cartilage and optimally capable of being resorbed into the body. Such matrices may be formed of materials presently in use for other implanted medical applications.

The choice of matrix material is based on biocompatibility, biodegradability, mechanical properties, cosmetic appearance and interface properties. The particular application of the compositions will define the appropriate formulation. Potential matrices for the compositions may be biodegradable and chemically defined calcium sulfate, tricalciumphosphate, hydroxyapatite, polylactic acid, polyglycolic acid and polyanhydrides. Other potential materials are biodegradable and biologically well-defined, such as bone or dermal collagen. Further matrices are comprised of pure proteins or extracellular matrix components. Other potential matrices are nonbiodegradable and chemically defined, such as sintered hydroxapatite, bioglass, aluminates, or other ceramics. Matrices may be comprised of combinations of any of the above mentioned types of material, such as polylactic acid and hydroxyapatite or collagen and tricalciumphosphate. The bioceramics may be altered in composition, such as in calcium-

aluminate-phosphate and processing to alter pore size, particle size, particle shape, and biodegradability.

Presently preferred is a 50:50 (mole weight) copolymer of lactic acid and glycolic acid in the form of porous particles having diameters ranging from 150 to 800 microns. In some applications, it will be useful to utilize a sequestering agent, such as carboxymethyl cellulose or autologous blood clot, to prevent the protein compositions from disassociating from the matrix.

A preferred family of sequestering agents is cellulosic materials such as alkylcelluloses (including hydroxyalkylcelluloses), including methylcellulose, ethylcellulose, hydroxyethylcellulose, hydroxypropylcellulose, hydroxypropyl-methylcellulose, and carboxymethylcellulose, the most preferred being cationic salts of carboxymethylcellulose (CMC). Other preferred sequestering agents include hyaluronic acid, sodium alginate, poly(ethylene glycol), polyoxyethylene oxide, carboxyvinyl polymer and poly(vinyl alcohol). The amount of sequestering agent useful herein is 0.5-20 wt%, preferably 1-10 wt% based on total formulation weight, which represents the amount necessary to prevent desorbtion of the protein from the polymer matrix and to provide appropriate handling of the composition, yet not so much that the progenitor cells are prevented from infiltrating the matrix, thereby providing the protein the opportunity to assist the osteogenic activity of the progenitor cells.

In further compositions, proteins of the invention may be combined with other agents beneficial to the treatment of the bone and/or cartilage defect, wound, or tissue in question. These agents include various growth factors such as epidermal growth factor (EGF), platelet derived growth factor (PDGF), transforming growth factors (TGF-$\alpha$ and TGF-$\beta$), and insulin-like growth factor (IGF).

The therapeutic compositions are also presently valuable for veterinary applications. Particularly domestic animals and thoroughbred horses, in addition to humans, are desired patients for such treatment with proteins of the present invention.

The dosage regimen of a protein-containing pharmaceutical composition to be used in tissue regeneration will be determined by the attending physician considering various factors which modify the action of the proteins, e.g., amount of tissue weight desired to be formed, the site of damage, the condition of the damaged tissue, the size of a wound, type of damaged tissue (e.g., bone), the patient's age, sex, and diet, the severity of any infection, time of administration and other clinical factors. The dosage may vary with the type of matrix used in the reconstitution and with inclusion of other proteins in

the pharmaceutical composition. For example, the addition of other known growth factors, such as IGF I (insulin like growth factor I), to the final composition, may also effect the dosage. Progress can be monitored by periodic assessment of tissue/bone growth and/or repair, for example, X-rays, histomorphometric determinations and tetracycline labeling.

Polynucleotides of the present invention can also be used for gene therapy. Such polynucleotides can be introduced either *in vivo* or *ex vivo* into cells for expression in a mammalian subject. Polynucleotides of the invention may also be administered by other known methods for introduction of nucleic acid into a cell or organism (including, without limitation, in the form of viral vectors or naked DNA).

Cells may also be cultured *ex vivo* in the presence of proteins of the present invention in order to proliferate or to produce a desired effect on or activity in such cells. Treated cells can then be introduced *in vivo* for therapeutic purposes.

Patent and literature references cited herein are incorporated by reference as if fully set forth.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Jacobs, Kenneth
McCoy, John M.
LaVallie, Edward R.
Racie, Lisa A.
Merberg, David
Treacy, Maurice
Spaulding, Vikki
Agostino, Michael

(ii) TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM

(iii) NUMBER OF SEQUENCES: 39

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Genetics Institute, Inc.
(B) STREET: 87 CambridgePark Drive
(C) CITY: Cambridge
(D) STATE: MA
(E) COUNTRY: U.S.A.
(F) ZIP: 02140

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Sprunger, Suzanne A.
(B) REGISTRATION NUMBER: 41,323

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (617) 498-8284
(B) TELEFAX: (617) 876-5851

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1605 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

ACGCTGACCA TCACAGGCAC ACAGAGGCAC ATCCACCTCA CATCCACCTC ATACTTGTGT 60

ACTCTCAGGG TTCAGTCTTT CATCCTATCC CTCTCTGATC TGTGCCTCCC AATACCTTCC 120

AAGATGTTTA CAGAGACCCT TCTCCCTGTG CAGTTAGGAG TGTAAGGCAA GAGAGCCCCT 180

ACTTCATGGG GCAGATCAAG AGCTGAGACC AAAGATGGTC TATGTTGCTG ACCTTGTCCT 240

GTCCTCCTGC TGTCTTAAAC TATGATCCCT GCTGCGGTCA CTGAAGCCTT TCCCTGTGAG 300

CAGTGGTGTG TGAGAGCCAG GCGTCCCTCT GCCTGCCCAC TCAGTGGCAA CACCCGGGAG 360

CTGTTTTGTC CTTTGTGGAG CCTCAGCAGT TCCCTCTTTC AGAACTCACT GCCAAGAGCC 420

CTGAACAGGA GCCACCATGC AGTGCTTCAG CTTCATTAAG ACCATGATGA TCCTCTTCAA 480

TTTGCTCATC TTTCTGTGTG GTGCAGCCCT GTTGGCAGTG GGCATCTGGG TGTCAATCGA 540

TGGGGCATCC TTTCTGAAGA TCTTCGGGCC ACTGTCGTCC AGTGCCATGC AGTTTGTCAA 600

CGTGGGCTAC TTCCTCATCG CAGCCGGCGT TGTGGTCTTT GCTCTTGGTT TCCTGGGCTG 660

CTATGGTGCT AAGACTGAGA GCAAGTGTGC CCTCGTGACG TTCTTCTTCA TCCTCCTCCT 720

CATCTTCATT GCTGAGGTTG CAGCTGCTGT GGTCGCCTTG GTGTACACCA CAATGGCTGA 780

GCACTTCCTG ACGTTGCTGG TAGTGCCTGC CATCAAGAAA GATTATGGTT CCCAGGAAGA 840

CTTCACTCAA GTGTGGAACA CCACCATGAA AGGGCTCAAG TGCTGTGGCT TCACCAACTA 900

TACGGATTTT GAGGACTCAC CCTACTTCAA AGAGAACAGT GCCTTTCCCC CATTCTGTTG 960

CAATGACAAC GTCACCAACA CAGCCAATGA AACCTGCACC AAGCAAAAGG CTCACGACCA 1020

AAAAGTAGAG GGTTGCTTCA ATCAGCTTTT GTATGACATC CGAACTAATG CAGTCACCGT 1080

GGGTGGTGTG GCAGCTGGAA TTGGGGGCCT CGAGCTGGCT GCCATGATTG TGTCCATGTA 1140

TCTGTACTGC AATCTACAAT AAGTCCACTT CTGCCTCTGC CACTACTGCT GCCACATGGG 1200

AACTGTGAAG AGGCACCCTG GCAAGCAGCA GTGATTGGGG GAGGGGACAG GATCTAACAA 1260

TGTCACTTGG GCCAGAATGG ACCTGCCCTT TCTGCTCCAG ACTTGGGGGT AGATAGGGAC 1320

CACTCCTTTT AGGCGATGCC TGACTTTCCT TCCATTGGTG GGTGGATGGG TGGGGGGCAT 1380

TCCAGAGCCT CTAAGGTAGC CAGTTCTGTT GCCCATTCCC CCAGTCTATT AAACCCTTGA 1440

TATGCCCCCT AGGCCTAGTG GTGATCCCAG TGCTCTACTG GGGGATGAGA GAAAGGCATT 1500

TTATAGCCTG GGCATAAGTG AAATCAGCAG AGCCTCTGGG TGGATGTGTA GAAGGCACTT 1560

CAAAATGCAT AAACCTGTTA CAATGTTGAA AAAAAAAAAA AAAAA 1605

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 241 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Gln Cys Phe Ser Phe Ile Lys Thr Met Met Ile Leu Phe Asn Leu
1 5 10 15

Leu Ile Phe Leu Cys Gly Ala Ala Leu Leu Ala Val Gly Ile Trp Val
20 25 30

Ser Ile Asp Gly Ala Ser Phe Leu Lys Ile Phe Gly Pro Leu Ser Ser
35 40 45

Ser Ala Met Gln Phe Val Asn Val Gly Tyr Phe Leu Ile Ala Ala Gly
50 55 60

Val Val Val Phe Ala Leu Gly Phe Leu Gly Cys Tyr Gly Ala Lys Thr
65 70 75 80

Glu Ser Lys Cys Ala Leu Val Thr Phe Phe Phe Ile Leu Leu Leu Ile
85 90 95

Phe Ile Ala Glu Val Ala Ala Ala Val Val Ala Leu Val Tyr Thr Thr
100 105 110

Met Ala Glu His Phe Leu Thr Leu Leu Val Val Pro Ala Ile Lys Lys
115 120 125

Asp Tyr Gly Ser Gln Glu Asp Phe Thr Gln Val Trp Asn Thr Thr Met
130 135 140

Lys Gly Leu Lys Cys Cys Gly Phe Thr Asn Tyr Thr Asp Phe Glu Asp
145 150 155 160

Ser Pro Tyr Phe Lys Glu Asn Ser Ala Phe Pro Pro Phe Cys Cys Asn
165 170 175

Asp Asn Val Thr Asn Thr Ala Asn Glu Thr Cys Thr Lys Gln Lys Ala
180 185 190

His Asp Gln Lys Val Glu Gly Cys Phe Asn Gln Leu Leu Tyr Asp Ile
195 200 205

Arg Thr Asn Ala Val Thr Val Gly Gly Val Ala Ala Gly Ile Gly Gly
210 215 220

Leu Glu Leu Ala Ala Met Ile Val Ser Met Tyr Leu Tyr Cys Asn Leu
225 230 235 240

Gln

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 377 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CAACCTCGTG GTCACCGCAC CGGGGCTGAT CAAGGGTGAC GCCTGCTTCA CATCTCTAAT 60

GAACACCCTC ATGACGTCGC TACCAGCACT AGTGCAGCAA CAGGGAAGGC TGCTTCTGGC 120

TGCTAATGTG GCCACCCTGG GGCTCCTCAT GGCCCGGCTC CTTAGCACCT CTCCAGCTCT 180

TCAGGGAACA CCAGCATCCC GAGGGTTCTT CGCAGCTGCC ATCCTCTTCC TATCACAGTC 240

CCACGTGGCG CGGGCCACCC CGGGCTCAGA CCAGGCAGTG CTAGCCCTGT CCCCTGAGTA 300

TGAGGGCATC TGGGCCGACC TGCAGGAGCT CTGGTTCCTG GGCATNCAAG CCTTCACCGG 360

CTGTGTGCCT CTGCTGC 377

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 106 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Asn Thr Leu Met Thr Ser Leu Pro Ala Leu Val Gln Gln Gln Gly
1 5 10 15

Arg Leu Leu Leu Ala Ala Asn Val Ala Thr Leu Gly Leu Leu Met Ala
20 25 30

Arg Leu Leu Ser Thr Ser Pro Ala Leu Gln Gly Thr Pro Ala Ser Arg
35 40 45

Gly Phe Phe Ala Ala Ala Ile Leu Phe Leu Ser Gln Ser His Val Ala
50 55 60

Arg Ala Thr Pro Gly Ser Asp Gln Ala Val Leu Ala Leu Ser Pro Glu
65 70 75 80

Tyr Glu Gly Ile Trp Ala Asp Leu Gln Glu Leu Trp Phe Leu Gly Xaa
85 90 95

Gln Ala Phe Thr Gly Cys Val Pro Leu Leu
100 105

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 245 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

TANGACTCCT CTCTGCNGAG ACGCGACTGG CGGNTCCAGC AGGGANTACC TTTTTTATAA 60

ACCCNGGGGG NCCACACACA CACACACACA CACACACACA CACACACACA CACACACACA 120

CATTTTTGAT CCCTTGCTTC CNTCCCCCAG TGCGTTCTGT GATCGCCAAG TTCAAAGCTG 180

TGCACATGTG GACACTCAAT AAATGTTCAT TGGNGACAAA AAAAAAAAAA AAAAAAAAAA 240

AAAAA 245

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2384 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

CCAAAACATG CTCAAAAGTA GAACATTTTG TTTCAATATT AGGAAAGTGC TTTGAATCCC 60

CTTGGACGAC AAAAGCGTTG TCTGAGACAG CATGCGAAGA CTCAGAGGAA AACAAGCAGA 120

GAATAACAGG TGCCCAGACT CTACCAAAGC ATGTTTCTAC CAGCAGTGAT GAAGGGAGCC 180

CCAGTGCCAG TACACCAATG ATCAATAAAA CTGGCTTTAA ATTTTCAGCT GAGAAGCCTG 240

TGATTGAAGT TCCCAGCATG ACAATCCTGG ATAAAAAGGA TGGAGAGCAG GCCAAAGCCC 300

TGTTTGAGAA AGTGAGGAAG TTCCGTGCCC ATGTGGAAGA TAGTGACTTG ATCTATAAAC 360

TCTATGTGGT CCAAACAGTT ATCAAAACAG CCAAGTTCAT TTTTATTCTC TGCTATACAG 420

CGAACTTTGT CAACGCAATC AGCTTTGAAC ACGTCTGCAA GCCCAAAGTT GAGCATCTGA 480

TTGGTTATGA GGTATTTGAG TGCACCCACA ATATGGCTTA CATGTTGAAA AAGCTTCTCA 540

TCAGTTACAT ATCCATTATT TGTGTTTATG GCTTTATCTG CCTCTACACT CTCTTCTGGT 600

TATTCAGGAT ACCTTTGAAG GAATATTCTT TCGAAAAAGT CAGAGAAGAG AGCAGTTTTA 660

GTGACATTCC AGATGTCAAA AACGATTTTG CGTTCCTTCT TCACATGGTA GACCAGTATG 720

ACCAGCTATA TTCCAAGCGT TTTGGTGTGT TCTTGTCAGA AGTTAGTGAA AATAAACTTA 780

GGGAAATTAG TTTGAACCAT GAGTGGACAT TTGAAAAACT CAGGCAGCAC ATTTCACGCA 840

ACGCCCAGGA CAAGCAGGAG TTGCATCTGT TCATGCTGTC GGGGGTGCCC GATGCTGTCT 900

TTGACCTCAC AGACCTGGAT GTGCTAAAGC TTGAACTAAT TCCAGAAGCT AAAATTCCTG 960

CTAAGATTTC TCAAATGACT AACCTCCAAG AGCTCCACCT CTGCCACTGC CCTGCAAAAG 1020

TTGAACAGAC TGCTTTTAGC TTTCTTCGCG ATCACTTGAG ATGCCTTCAC GTGAAGTTCA 1080

CTGATGTGGC TGAAATTCCT GCCTGGGTGT ATTTGCTCAA AAACCTTCGA GAGTTGTACT 1140

TAATAGGCAA TTTGAACTCT GAAAACAATA AGATGATAGG ACTTGAATCT CTCCGAGAGT 1200

TGCGGCACCT TAAGATTCTC CACGTGAAGA GCAATTTGAC CAAAGTTCCC TCCAACATTA 1260

CAGATGTGGC TCCACATCTT ACAAAGTTAG TCATTCATAA TGACGGCACT AAACTCTTGG 1320

TACTGAACAG CCTTAAGAAA ATGATGAATG TCGCTGAGCT GGAACTCCAG AACTGTGAGC 1380

TAGAGAGAAT CCCACATGCT ATTTTCAGCC TCTCTAATTT ACAGGAACTG GATTTAAAGT 1440

CCAATAACAT TCGCACAATT GAGGAAATCA TCAGTTTCCA GCATTTAAAA CGACTGACTT 1500

GTTTAAAATT ATGGCATAAC AAAATTGTTA CTATTCCTCC CTCTATTACC CATGTCAAAA 1560

ACTTGGAGTC ACTTTATTTC TCTAACAACA AGCTCGAATC CTTACCAGTG GCAGTATTTA 1620

GTTTACAGAA ACTCAGATGC TTAGATGTGA GCTACAACAA CATTTCAATG ATTCCAATAG 1680

AAATAGGATT GCTTCAGAAC CTGCAGCATT TGCATATCAC TGGGAACAAA GTGGACATTC 1740

TGCCAAAACA ATTGTTTAAA TGCATAAAGT TGAGGACTTT GAATCTGGGA CAGAACTGCA 1800

TCACCTCACT CCCAGAGAAA GTTGGTCAGC TCTCCCAGCT CACTCAGCTG GAGCTGAAGG 1860

GGAACTGCTT GGACCGCCTG CCAGCCCAGC TGGGCCAGTG TCGGATGCTC AAGAAAAGCG 1920

GGCTTGTTGT GGAAGATCAC CTTTTTGATA CCCTGCCACT CGAAGTCAAA GAGGCATTGA 1980

ATCAAGACAT AAATATTCCC TTTGCAAATG GGATTTAMAC TAAGATAATA TATGCACAGT 2040

GATGTGCAGG AACAACTTCC TAGATTGCAA GTGCTCACGT ACAAGTTATT ACAAGATAAT 2100

GCATTTTAGG AGTAGATACA TCTTTTAAAA TAAAACAGAG AGGATGCATA GAAGGCTGAT 2160

AGAAGACATA ACTGAATGTT CAATGTTTGT AGGGTTTTAA GTCATTCATT TCCAAATCAT 2220

TTTTTTTTTT CTTTTGGGGA AAGGGAAGGA AAAATTATAA TCACTAATCT TGGTTCTTTT 2280

TAAATTGTTT GTAACTTGGA TGCTGCCGCT ACTGAATGTT TACAAATTGC TTGCCTGCTA 2340

AAGTAAATGA TTAAATTGAC ATTTTCTTAC TATAAAAAAA AAAA 2384

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 614 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

```
Met Ile Asn Lys Thr Gly Phe Lys Phe Ser Ala Glu Lys Pro Val Ile
1               5                  10                  15

Glu Val Pro Ser Met Thr Ile Leu Asp Lys Lys Asp Gly Glu Gln Ala
            20                  25                  30

Lys Ala Leu Phe Glu Lys Val Arg Lys Phe Arg Ala His Val Glu Asp
        35                  40                  45

Ser Asp Leu Ile Tyr Lys Leu Tyr Val Val Gln Thr Val Ile Lys Thr
    50                  55                  60

Ala Lys Phe Ile Phe Ile Leu Cys Tyr Thr Ala Asn Phe Val Asn Ala
65                  70                  75                  80
```

```
Ile Ser Phe Glu His Val Cys Lys Pro Lys Val Glu His Leu Ile Gly
                85                  90                  95

Tyr Glu Val Phe Glu Cys Thr His Asn Met Ala Tyr Met Leu Lys Lys
            100                 105                 110

Leu Leu Ile Ser Tyr Ile Ser Ile Ile Cys Val Tyr Gly Phe Ile Cys
        115                 120                 125

Leu Tyr Thr Leu Phe Trp Leu Phe Arg Ile Pro Leu Lys Glu Tyr Ser
    130                 135                 140

Phe Glu Lys Val Arg Glu Glu Ser Ser Phe Ser Asp Ile Pro Asp Val
145                 150                 155                 160

Lys Asn Asp Phe Ala Phe Leu Leu His Met Val Asp Gln Tyr Asp Gln
                165                 170                 175

Leu Tyr Ser Lys Arg Phe Gly Val Phe Leu Ser Glu Val Ser Glu Asn
            180                 185                 190

Lys Leu Arg Glu Ile Ser Leu Asn His Glu Trp Thr Phe Glu Lys Leu
        195                 200                 205

Arg Gln His Ile Ser Arg Asn Ala Gln Asp Lys Gln Glu Leu His Leu
    210                 215                 220

Phe Met Leu Ser Gly Val Pro Asp Ala Val Phe Asp Leu Thr Asp Leu
225                 230                 235                 240

Asp Val Leu Lys Leu Glu Leu Ile Pro Glu Ala Lys Ile Pro Ala Lys
                245                 250                 255

Ile Ser Gln Met Thr Asn Leu Gln Glu Leu His Leu Cys His Cys Pro
            260                 265                 270

Ala Lys Val Glu Gln Thr Ala Phe Ser Phe Leu Arg Asp His Leu Arg
        275                 280                 285

Cys Leu His Val Lys Phe Thr Asp Val Ala Glu Ile Pro Ala Trp Val
    290                 295                 300

Tyr Leu Leu Lys Asn Leu Arg Glu Leu Tyr Leu Ile Gly Asn Leu Asn
305                 310                 315                 320

Ser Glu Asn Asn Lys Met Ile Gly Leu Glu Ser Leu Arg Glu Leu Arg
                325                 330                 335

His Leu Lys Ile Leu His Val Lys Ser Asn Leu Thr Lys Val Pro Ser
            340                 345                 350

Asn Ile Thr Asp Val Ala Pro His Leu Thr Lys Leu Val Ile His Asn
        355                 360                 365

Asp Gly Thr Lys Leu Leu Val Leu Asn Ser Leu Lys Lys Met Met Asn
    370                 375                 380
```

```
Val Ala Glu Leu Glu Leu Gln Asn Cys Glu Leu Glu Arg Ile Pro His
385                 390                 395                 400

Ala Ile Phe Ser Leu Ser Asn Leu Gln Glu Leu Asp Leu Lys Ser Asn
                405                 410                 415

Asn Ile Arg Thr Ile Glu Glu Ile Ile Ser Phe Gln His Leu Lys Arg
            420                 425                 430

Leu Thr Cys Leu Lys Leu Trp His Asn Lys Ile Val Thr Ile Pro Pro
        435                 440                 445

Ser Ile Thr His Val Lys Asn Leu Glu Ser Leu Tyr Phe Ser Asn Asn
    450                 455                 460

Lys Leu Glu Ser Leu Pro Val Ala Val Phe Ser Leu Gln Lys Leu Arg
465                 470                 475                 480

Cys Leu Asp Val Ser Tyr Asn Asn Ile Ser Met Ile Pro Ile Glu Ile
                485                 490                 495

Gly Leu Leu Gln Asn Leu Gln His Leu His Ile Thr Gly Asn Lys Val
            500                 505                 510

Asp Ile Leu Pro Lys Gln Leu Phe Lys Cys Ile Lys Leu Arg Thr Leu
        515                 520                 525

Asn Leu Gly Gln Asn Cys Ile Thr Ser Leu Pro Glu Lys Val Gly Gln
    530                 535                 540

Leu Ser Gln Leu Thr Gln Leu Glu Leu Lys Gly Asn Cys Leu Asp Arg
545                 550                 555                 560

Leu Pro Ala Gln Leu Gly Gln Cys Arg Met Leu Lys Lys Ser Gly Leu
                565                 570                 575

Val Val Glu Asp His Leu Phe Asp Thr Leu Pro Leu Glu Val Lys Glu
            580                 585                 590

Ala Leu Asn Gln Asp Ile Asn Ile Pro Phe Ala Asn Gly Ile Xaa Thr
        595                 600                 605

Lys Ile Ile Tyr Ala Gln
    610
```

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2386 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

CCAAAACATC AAACCCTTTT CTTGTAGCAG CACAGGATTC TGAGACAGAT TATGTCACAA 60

CAGATAATTT AACAAAGGTG ACTGAGGAAG TCGTGGCAAA CATGCCTGAA GGCCTGACTC 120

CAGATTTAGT ACAGGAAGCA TGTGAAAGTG AATTGAATGA AGTTACTGGT ACAAAGATTG 180

CTTATGAAAC AAAAATGGAC TTGGTTCAAA CATCAGAAGT TATGCAAGAG TCACTCTATC 240

CTGCAGCACA GCTTTGCCCA TCATTTGAAG AGTCAGAAGC TACTCCTTCA CCAGTTTTGC 300

CTGACATTGT TATGGAAGCA CCATTGAATT CTGCAGTTCC TAGTGCTGGT GCTTCCGTGA 360

TACAGCCCAG CTCATCACCA TTAGAAGCTT CTTCAGTTAA TTATGAAAGC ATAAAACATG 420

AGCCTGAAAA CCCCCCACCA TATGAAGAGG CCATGAGTGT ATCACTAAAA AAAGTATCAG 480

GAATAAAGGA AGAAATTAAA GAGCCTGAAA ATATTAATGC AGCTCTTCAA GAAACAGAAG 540

CTCCTTATAT ATCTATTGCA TGTGATTTAA TTAAAGAAAC AAAGCTTTCT GCTGAACCAG 600

CTCCGGATTT CTCTGATTAT TCAGAAATGG CAAAAGTTGA ACAGCCAGTG CCTGATCATT 660

CTGAGCTAGT TGAAGATTCC TCACCTGATT CTGAACCAGT TGACTTATTT AGTGATGATT 720

CAATACCTGA CGTTCCACAA AAACAAGATG AAACTGTGAT GCTTGTGAAA GAAAGTCTCA 780

CTGAGACTTC ATTTGAGTCA ATGATAGAAT ATGAAAATAA GGAAAAACTC AGTGCTTTGC 840

CACCTGAGGG AGGAAAGCCA TATTTGGAAT CTTTTAAGCT CAGTTTAGAT AACACAAAAG 900

ATACCCTGTT ACCTGATGAA GTTTCAACAT TGAGCAAAAA GGAGAAAATT CCTTTGCAGA 960

TGGAGGAGCT CAGTACTGCA GTTTATTCAA ATGATGACTT ATTTATTTCT AAGGAAGCAC 1020

AGATAAGAGA AACTGAAACG TTTTCAGATT CATCTCCAAT TGAAATTATA GATGAGTTCC 1080

CTACATTGAT CAGTTCTAAA ACTGATTCAT TTTCTAAATT AGCCAGGGAA TATACTGACC 1140

TAGAAGTATC CCACAAAAGT GAAATTGCTA ATGCCCCGGA TGGAGCTGGG TCATTGCCTT 1200

GCACAGAATT GCCCCATGAC CTTTCTTTGA AGAACATACA ACCCAAAGTT GAAGAGAAAA 1260

TCAGTTTCTC AGATGACTTT TCTAAAAATG GGTCTGCTAC ATCAAAGGTG CTCTTATTGC 1320

CTCCAGATGT TTCTGCTTTG GCCACTCAAG CAGAGATAGA GAGCATAGTT AAACCCAAAG 1380

TTCTTGTGAA AGAAGCTGAG AAAAAACTTC CTTCCGATAC AGAAAAAGAG GACAGATCAC 1440

CATCTGCTAT ATTTTCAGCA GAGCTGAGTA AAACTTCAGT TGTTGACCTC CTGTACTGGA 1500

GAGACATTAA GAAGACTGGA GTGGTGTTTG GTGCCAGCCT ATTCCTGCTG CTTTCATTGA 1560

CAGTATTCAG CATTGTGAGC GTAACAGCCT ACATTGCCTT GGCCCTGCTC TCTGTGACCA 1620

TCAGCTTTAG GATATACAAG GGTGTGATCC AAGCTATCCA GAAATCAGAT GAAGGCCACC 1680

CATTCAGGGA AGTTGCTATA TCTGAGGAGT TGGTTCAGAA GTACAGTAAT TCTGCTCTTG 1740

GTCATGTGAA CTGCACGATA AAGGAACTCA GGCGCCTCTT CTTAGTTGAT GATTTAGTTG 1800

ATTCTCTGAA GTTTGCAGTG TTGATGTGGG TATTTACCTA TGTTGGTGCC TTGTTTAATG 1860

GTCTGACACT ACTGATTTTG GCTCTCATTT CACTCTTCAG TGTTCCTGTT ATTTATGAAC 1920

GGCATCAGGC ACAGATAGAT CATTATCTAG GACTTGCAAA TAAGAATGTT AAAGATGCTA 1980

TGGCTAAAAT CCAAGCAAAA ATCCCTGGAT TGAAGCGCAA AGCTGAATGA AAACGCCCAA 2040

AATAATTAGT AGGAGTTCAT CTTTAAAGGG GATATTCATT TGATTATACG GGGGAGGGTC 2100

AGGGAAGAAC GAACCTTGAC GTTGCAGTGC AGTTTCACAG ATCGTTGTTA GATCTTTATT 2160

TTTAGCCATG CACTGTTGTG AGGAAAAATT ACCTGTCTTG ACTGCCATGT GTTCATCATC 2220

TTAAGTATTG TAAGCTGCTA TGTATGGATT TAAACCGTAA TCATATCTTT TTCCTATCTG 2280

AGGCACTGGT GGAATAAAAA ACCTGTATAT TTTACTTTGT TGCAGATAGT CTTGCCGCAT 2340

CTTGGCAAGT TGCAGAGATG GTGGAGCTAG AAAAAAAAAA AAAAAA 2386

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 642 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Met Pro Glu Gly Leu Thr Pro Asp Leu Val Gln Glu Ala Cys Glu Ser
1 5 10 15

Glu Leu Asn Glu Val Thr Gly Thr Lys Ile Ala Tyr Glu Thr Lys Met
20 25 30

Asp Leu Val Gln Thr Ser Glu Val Met Gln Glu Ser Leu Tyr Pro Ala
35 40 45

Ala Gln Leu Cys Pro Ser Phe Glu Glu Ser Glu Ala Thr Pro Ser Pro
50 55 60

Val Leu Pro Asp Ile Val Met Glu Ala Pro Leu Asn Ser Ala Val Pro
65 70 75 80

```
Ser Ala Gly Ala Ser Val Ile Gln Pro Ser Ser Ser Pro Leu Glu Ala
                85                  90                  95

Ser Ser Val Asn Tyr Glu Ser Ile Lys His Glu Pro Glu Asn Pro Pro
            100                 105                 110

Pro Tyr Glu Glu Ala Met Ser Val Ser Leu Lys Lys Val Ser Gly Ile
        115                 120                 125

Lys Glu Glu Ile Lys Glu Pro Glu Asn Ile Asn Ala Ala Leu Gln Glu
    130                 135                 140

Thr Glu Ala Pro Tyr Ile Ser Ile Ala Cys Asp Leu Ile Lys Glu Thr
145                 150                 155                 160

Lys Leu Ser Ala Glu Pro Ala Pro Asp Phe Ser Asp Tyr Ser Glu Met
                165                 170                 175

Ala Lys Val Glu Gln Pro Val Pro Asp His Ser Glu Leu Val Glu Asp
            180                 185                 190

Ser Ser Pro Asp Ser Glu Pro Val Asp Leu Phe Ser Asp Asp Ser Ile
        195                 200                 205

Pro Asp Val Pro Gln Lys Gln Asp Glu Thr Val Met Leu Val Lys Glu
    210                 215                 220

Ser Leu Thr Glu Thr Ser Phe Glu Ser Met Ile Glu Tyr Glu Asn Lys
225                 230                 235                 240

Glu Lys Leu Ser Ala Leu Pro Pro Glu Gly Gly Lys Pro Tyr Leu Glu
                245                 250                 255

Ser Phe Lys Leu Ser Leu Asp Asn Thr Lys Asp Thr Leu Leu Pro Asp
            260                 265                 270

Glu Val Ser Thr Leu Ser Lys Lys Glu Lys Ile Pro Leu Gln Met Glu
        275                 280                 285

Glu Leu Ser Thr Ala Val Tyr Ser Asn Asp Asp Leu Phe Ile Ser Lys
    290                 295                 300

Glu Ala Gln Ile Arg Glu Thr Glu Thr Phe Ser Asp Ser Ser Pro Ile
305                 310                 315                 320

Glu Ile Ile Asp Glu Phe Pro Thr Leu Ile Ser Ser Lys Thr Asp Ser
                325                 330                 335

Phe Ser Lys Leu Ala Arg Glu Tyr Thr Asp Leu Glu Val Ser His Lys
            340                 345                 350

Ser Glu Ile Ala Asn Ala Pro Asp Gly Ala Gly Ser Leu Pro Cys Thr
        355                 360                 365

Glu Leu Pro His Asp Leu Ser Leu Lys Asn Ile Gln Pro Lys Val Glu
    370                 375                 380
```

```
Glu Lys Ile Ser Phe Ser Asp Asp Phe Ser Lys Asn Gly Ser Ala Thr
385                 390                 395                 400

Ser Lys Val Leu Leu Leu Pro Pro Asp Val Ser Ala Leu Ala Thr Gln
                405                 410                 415

Ala Glu Ile Glu Ser Ile Val Lys Pro Lys Val Leu Val Lys Glu Ala
            420                 425                 430

Glu Lys Lys Leu Pro Ser Asp Thr Glu Lys Glu Asp Arg Ser Pro Ser
        435                 440                 445

Ala Ile Phe Ser Ala Glu Leu Ser Lys Thr Ser Val Val Asp Leu Leu
    450                 455                 460

Tyr Trp Arg Asp Ile Lys Lys Thr Gly Val Val Phe Gly Ala Ser Leu
465                 470                 475                 480

Phe Leu Leu Leu Ser Leu Thr Val Phe Ser Ile Val Ser Val Thr Ala
                485                 490                 495

Tyr Ile Ala Leu Ala Leu Leu Ser Val Thr Ile Ser Phe Arg Ile Tyr
            500                 505                 510

Lys Gly Val Ile Gln Ala Ile Gln Lys Ser Asp Glu Gly His Pro Phe
        515                 520                 525

Arg Glu Val Ala Ile Ser Glu Glu Leu Val Gln Lys Tyr Ser Asn Ser
    530                 535                 540

Ala Leu Gly His Val Asn Cys Thr Ile Lys Glu Leu Arg Arg Leu Phe
545                 550                 555                 560

Leu Val Asp Asp Leu Val Asp Ser Leu Lys Phe Ala Val Leu Met Trp
                565                 570                 575

Val Phe Thr Tyr Val Gly Ala Leu Phe Asn Gly Leu Thr Leu Leu Ile
            580                 585                 590

Leu Ala Leu Ile Ser Leu Phe Ser Val Pro Val Ile Tyr Glu Arg His
        595                 600                 605

Gln Ala Gln Ile Asp His Tyr Leu Gly Leu Ala Asn Lys Asn Val Lys
    610                 615                 620

Asp Ala Met Ala Lys Ile Gln Ala Lys Ile Pro Gly Leu Lys Arg Lys
625                 630                 635                 640

Ala Glu
```

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 344 base pairs
(B) TYPE: nucleic acid

(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

GGCGGCTGCG GANCCGGCGG TCCTTGCGCT CCCCAACANC GGCGCCGGGG GCGCGGGGGC 60

GCCGTCGGGC ACAGTCCCGG TGCTCTTCTG TTTCTCAGTC TTCGCGCGAC CCTCGTCGGT 120

GCCACACGGG GCGGGCTACA AGCTGCTCAT CCAGAAGTTC CTCAGCCTGT ACGGCGACCA 180

GATCNACATG CACCGCAAAT TCGTGGTGCA GCTGTTCGCC GAGGAGTGGG GCCAGTACGT 240

GGACTTGCCC AAGGGCTTCN CGGTGAGCGA GCGCTGCAAG GTGCGCCTCG TGCCGCTGCA 300

TATCCAGCTC ACTACCCTGG GAAATCTTAC ACCTTCAAGC ACTG 344

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 631 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

ACATGATTCA GAGTCCTGTG GGTAAATTCA TATGCAATAA TCTTATTCCA ATCAATCTGT 60

AAAGTAAAAG CANTACATCC ACATTAACAT TATAACATCT TACAGTAATA TAAAAGCCAA 120

ATCATTGTTG GTACGTCATT TTCTTTAAAG TGAACAATTT AAGAAAACTT CACAAGAGTC 180

TGCACTTTGG AAAGATACGA TCAGAGTACA CAGTAGAGAC AAAACAGGCA TCTTCATTGT 240

AATTTTTTTT AATAAATAAA AGCACATTAA CAAAAAAGGA AGGTAAGCAG CACCGGAAGC 300

CTTTGACGTT TGTAACTAAA TGCTGGTACT CAATTGAATC GAGCTGGTTA AGTTTCACTA 360

GGAGGCGCNA AAAAGGAGCC GTTTTTGACT TAACATTTTA ATTCTAGTAG AGATAAGAAG 420

AGCTTGTGTG GGCTTACAGT CCTTCACCTG ACTGTCCTTC ACCAGTGAGT AGCATACCAG 480

TTCTTCAAAT GTCCTATACT TTGGAAAGCA GACCCGACTC TGGAGCACTC GCCTTAATTA 540

GATTCTGAAT TTCCTTGAAT TTTGGATGGT CCTTATCAGC TACCAGCTGA AGCAGAACAG 600

CCTCACTCGT GGTCACTATG ATCCCGGTTC G 631

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Met Val Leu Ile Ser Tyr Gln Leu Lys Gln Asn Ser Leu Thr Arg Gly
1 5 10 15

His Tyr Asp Pro Gly Ser
20

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 70 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA 60

AAAAAAAAAA 70

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 428 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

AGCACGCGGT CCTGCCCGTG GACGGGGCAA CGCTGGCAGA TGTGATGCGC CAGCGGGGCA 60

TCAACATGCG CTACCTGGGC AAGGTGCTGG AGCTGGTGCT GCGGARCCCG GCCCGCCACC 120

AGCTGGACCA CGTCTTTAAA ATCGGCATTG GAGAACTCAT CACCCGCTCG SCCAAGCACA 180

TCTTCAAGAC GTACTTACAG GGAGTCGAGC TCTCCGGCCT CTCAGCCGCC ATCAGCCACT 240

TCCTGAACTG CTTCCTGAGC TCCTACCCAA ACCCCGTGGC CCACCTGCCC GCCGACGAGC 300

TGGTCTCCAA GAAGCGGAAT AAGAGGAGGA AAAACCGGCC CCCGGGGGCT GCAGATAACA 360

CAGCCTGGGC TGTCATGACC CCCCAGGAGC TCTGGAAGAA CATCTGCCAG GAGGCCAAGA 420

ACTACTTT 428

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 128 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

Met Arg Gln Arg Gly Ile Asn Met Arg Tyr Leu Gly Lys Val Leu Glu
1 5 10 15

Leu Val Leu Arg Xaa Pro Ala Arg His Gln Leu Asp His Val Phe Lys
20 25 30

Ile Gly Ile Gly Glu Leu Ile Thr Arg Ser Xaa Lys His Ile Phe Lys
35 40 45

Thr Tyr Leu Gln Gly Val Glu Leu Ser Gly Leu Ser Ala Ala Ile Ser
50 55 60

His Phe Leu Asn Cys Phe Leu Ser Ser Tyr Pro Asn Pro Val Ala His
65 70 75 80

Leu Pro Ala Asp Glu Leu Val Ser Lys Lys Arg Asn Lys Arg Arg Lys
85 90 95

Asn Arg Pro Pro Gly Ala Ala Asp Asn Thr Ala Trp Ala Val Met Thr
100 105 110

Pro Gln Glu Leu Trp Lys Asn Ile Cys Gln Glu Ala Lys Asn Tyr Phe
115 120 125

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 245 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

TGGTGGGGGA GGATGTGCCC ACCCTGAGAC CCGGAGGAGA CGGGCNTTTG CCTGGGTTTG 60

CGGAGAGCCG CTTATGGGTG TGGTCCGTCC AGACACCTTG TTTCAAGGGG GATGGGCGTG 120

AGCGGGCAAG CAGAGCANCC CCACCGNTGA GCAAGAACTT TTTTTTGTTT TTAAACCATC 180

ACGTCCTCAT TTCACATTGG AATAAAGTGA GTTTTTGAAA AAAAAAAAAA AAAAAAAAAA 240

AAAAA 245

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 566 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

CAGTGAGCCC TTTGAAAAAT AAACATCCAG ATGAAGATGC TGTGGAAGCT GAGGGGCATG 60

AGGTAAAAAG ACTCAGGTTT GACAAAGAAG GTGAAGTCAG AGAAACAGCC AGTCAAACGA 120

CTTCCAGCGA AATTTCTTCA GTTATGGTAG GAGAAACAGA AGCATCATCT TCATCTCAGG 180

ATAAAGACAA AGATAGCCGT TGTWCCCGGC AGCACTGTWC AGAAGAGGAT GAAGAAGAGG 240

ATGAAGAGGA AGAAGAAGAG TCTTTTATGA CATCAAGAGA AATGATCCCA GAAAGAAAAA 300

ATCAAGAAAA AGAATCTGAT GATGCCTTAA CTGTGAATGA AGAGACTTCT GAGGAAAATA 360

ATCAAATGGA GGAATCTGAT GTGTCTCAAG CTGAGAAAGA TTTGCTACAT TCTGAAGGTA 420

GTGAAAACGA AGGCCCTGTA AGTAGTAGTT CTTCTGACTG CCGTGAAACA GAAGAATTAG 480

TAGGATCCAA TTCCAGTAAA ACTGGAGAGA TTCTTTCAGA ATCATCCATG GAAAATGATG 540

ACGAAGCCAC AGAAGTCACC GATGAA 566

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 141 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

Met Val Gly Glu Thr Glu Ala Ser Ser Ser Ser Gln Asp Lys Asp Lys
1 5 10 15

Asp Ser Arg Cys Xaa Arg Gln His Cys Xaa Glu Glu Asp Glu Glu Glu
20 25 30

Asp Glu Glu Glu Glu Glu Glu Ser Phe Met Thr Ser Arg Glu Met Ile
35 40 45

Pro Glu Arg Lys Asn Gln Glu Lys Glu Ser Asp Asp Ala Leu Thr Val
50 55 60

Asn Glu Glu Thr Ser Glu Glu Asn Asn Gln Met Glu Glu Ser Asp Val
65 70 75 80

Ser Gln Ala Glu Lys Asp Leu Leu His Ser Glu Gly Ser Glu Asn Glu
85 90 95

Gly Pro Val Ser Ser Ser Ser Ser Asp Cys Arg Glu Thr Glu Glu Leu
100 105 110

Val Gly Ser Asn Ser Ser Lys Thr Gly Glu Ile Leu Ser Glu Ser Ser
115 120 125

Met Glu Asn Asp Asp Glu Ala Thr Glu Val Thr Asp Glu
130 135 140

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 531 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

TCATCATGGC TATAAATACC AAAACGATTT GGATCCATTT ATGTTTGTAG GATAATATAC 60

TACTGACTGA CTTGACTGTC AGGTTCACAA CAGCTAGATG ATATATTTAT GACTATGTCT 120

AATAGTTGAA ATAAAATCTG AATATTGATT TACTATACCC AAGAGGGGAG AAAAATTAAC 180

CATTGTAAAT TTTTAAAAAT TTTTTCAAAA ATGTTAAAAT GAGGCAAATT TAAGTTTACA 240

AATTTTGAAA TTTTCTTTTG AATATTTATG AAATTGTCAG TAAACTTACC TAAGATCCTG 300

TGACCTTTTG ATATTTTTTA TTTTAATTGT AGTGCCATGG ACCATTTGTA AACAAATTGA 360

TTTACTTTTG TTGGTTGTAA GTTGAAGATT TAGCATTATG ACTTTGAGGT CTGTGGTTTT 420

ATTTGTAAAC TTGCAATTGC TATATTTGCA AGGGCAAATG TATTTCTTTA TTAAATAAAG 480

TACAATAATG GTGAATGTAC CAAAATGACA TCACTTAAAA AAAAAAAAAA A 531

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1163 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

KCTGGAACCA CGCGGARGAA GGAAGAGACG CAGGCAGGCT GCGGTTACCC AAGCGGSCAC 60

CCGGGCCTCA GGGACCCTTC CCCGAGAGAC GGCACCATGA CCCAGGGAAA GCTCTCCGTG 120

GCTAACAAGC CCCTGGGACC GAGGGGCAGC AGCAKGTGCA TGGCGAGAAG AAGGAGCTCC 180

AGCAGTGCCC TCAGCCCCAC CCTCCTATGA GGAACCACCT CTGGGGAGGG GATGAAGGCA 240

GGGGCCTTCC CCCCAGCCCC CACAGCGGTG CCTCTCCACC CTAGCTGGGC CTATGTGGAC 300

CCCAGCAGCA GCTCCAGCTA TGACAACGGT TTCCCCACCG GAGACCATGA GCTCTTCACC 360

ACTTTCAGCT GGGATGACCA GAAAGTTCGT CGAGTCTTTG TCAGAAAGGT CTACACCATC 420

CTGCTGATTC AGCTGCTGGT GACCTTGGCT GTCGTGGCTC TCTTTACTTT CTGTGACCCT 480

GTCAAGGACT ATGTCCAGGC CAACCCAGGC TGGTACTGGG CATCCTATGC TGTGTTCTTT 540

GCAACCTACC TGACCCTGGC TTGCTGTTCT GGACCCAGGA GGCATTTCCC CTGGAACCTG 600

ATTCTCCTGA CCGTCTTTAC CCTGTCCATG GCCTACCTCA CTGGGATGCT GTCCAGCTAC 660

TACAACACCA CCTCCGTGCT GCTGTGCCTG GGCATCACGG CCCTTGTCTG CCTCTCAGTC 720

ACCGTCTTCA GCTTCCAGAC CAAGTTCGAC TTCACCTCCT GCCAGGGCGT GCTCTTCGTG 780

CTTCTCATGA CTCTTTTCTT CAGCGGACTC ATCCTGGCCA TCCTCCTACC CTTCCAATAT 840

GTGCCCTGGC TCCATGCAGT TTATGCAGCA CTGGGAGCGG GTGTATTTAC ATTGTTCCTG 900

GCACTTGACA CCCAGTTGCT GATGGGTAAC CGACGCCACT CGCTGAGCCC TGAGGAGTAT 960

ATTTTTGGAG CCCTCAACAT TTACCTAGAC ATCATCTATA TCTTCACCTT CTTCCTGCAG 1020

CTTTTTGGCA CTAACCGAGA ATGAGGAGCC CTCCCTGCCC CACCGTCCTC CAGAGAATGC 1080

GCCCCTCCTG GTTCCCTGTC CCTCCCCTGC GCTCCTGCGA GACCAGATAT AAAACTAGCT 1140

GCCAACCCAA AAAAAAAAAA AAA 1163

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 270 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

Met Lys Ala Gly Ala Phe Pro Pro Ala Pro Thr Ala Val Pro Leu His
1 5 10 15

Pro Ser Trp Ala Tyr Val Asp Pro Ser Ser Ser Ser Ser Tyr Asp Asn
20 25 30

Gly Phe Pro Thr Gly Asp His Glu Leu Phe Thr Thr Phe Ser Trp Asp
35 40 45

Asp Gln Lys Val Arg Arg Val Phe Val Arg Lys Val Tyr Thr Ile Leu
50 55 60

Leu Ile Gln Leu Leu Val Thr Leu Ala Val Val Ala Leu Phe Thr Phe
65 70 75 80

Cys Asp Pro Val Lys Asp Tyr Val Gln Ala Asn Pro Gly Trp Tyr Trp
85 90 95

Ala Ser Tyr Ala Val Phe Phe Ala Thr Tyr Leu Thr Leu Ala Cys Cys
100 105 110

```
Ser Gly Pro Arg Arg His Phe Pro Trp Asn Leu Ile Leu Leu Thr Val
        115                 120                 125

Phe Thr Leu Ser Met Ala Tyr Leu Thr Gly Met Leu Ser Ser Tyr Tyr
    130                 135                 140

Asn Thr Thr Ser Val Leu Leu Cys Leu Gly Ile Thr Ala Leu Val Cys
145                 150                 155                 160

Leu Ser Val Thr Val Phe Ser Phe Gln Thr Lys Phe Asp Phe Thr Ser
                165                 170                 175

Cys Gln Gly Val Leu Phe Val Leu Leu Met Thr Leu Phe Phe Ser Gly
            180                 185                 190

Leu Ile Leu Ala Ile Leu Leu Pro Phe Gln Tyr Val Pro Trp Leu His
        195                 200                 205

Ala Val Tyr Ala Ala Leu Gly Ala Gly Val Phe Thr Leu Phe Leu Ala
    210                 215                 220

Leu Asp Thr Gln Leu Leu Met Gly Asn Arg Arg His Ser Leu Ser Pro
225                 230                 235                 240

Glu Glu Tyr Ile Phe Gly Ala Leu Asn Ile Tyr Leu Asp Ile Ile Tyr
                245                 250                 255

Ile Phe Thr Phe Phe Leu Gln Leu Phe Gly Thr Asn Arg Glu
            260                 265                 270
```

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 624 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

```
CGCACCCCCT CCGGCCGCGG GCGCAKCGGG GGCGCTGGTG GAKCTGMGAA GGGCCASGTC     60

CGGCGGGCGG GGCGGCGGCT GGCACTGGCT CCGGACTCTG CCCGGCCAGG GCGGCGGMTC    120

CANCCGGGAG GGCGACGTGG AGCGGCCACK TGGAKCGGCC CGGGGGARGC TGGCGGCGGG    180

AKGCGAGGCG CGGGCGGCGC AKCAKCCAKG AGCGCCCACG GAGSTGGACC CCCAGAKCCG    240

CGCGGCGCCG CAGCAGTTCC AGGAAGGATG TTACCTTTGA CGATGACAGT GTTAATCCTG    300

CTGCTGCTCC CCACGGGTCA GGCTGCCCCA AAGGATGGAG TCACAAGGCC AGAATCTGAA    360
```

GTGCAGCATC AGCTCCTGCC CAACCCCTTC CAGCCAGGCC AGGAGCAGCT CGGACTTCTG 420

CAGAGCTACC TAAAGGGACT AGGAAGGACA GAAGTGCAAC TGGAGCATCT GAGCCGGGAG 480

CAGGTTCTCC TCTACCTCTT TGCCCTCCAT GACTATGACC AGAGTGGACA GCTGGATGGC 540

CTGGAGCTGC TGTCCATGTT GACAGCTGCT CTGGCCCCTG GAGCTGCCAA CTCTCCTACC 600

ACCAACCCGG TGATCTTGAT AGTG 624

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 119 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

Met Leu Pro Leu Thr Met Thr Val Leu Ile Leu Leu Leu Leu Pro Thr
1 5 10 15

Gly Gln Ala Ala Pro Lys Asp Gly Val Thr Arg Pro Glu Ser Glu Val
20 25 30

Gln His Gln Leu Leu Pro Asn Pro Phe Gln Pro Gly Gln Glu Gln Leu
35 40 45

Gly Leu Leu Gln Ser Tyr Leu Lys Gly Leu Gly Arg Thr Glu Val Gln
50 55 60

Leu Glu His Leu Ser Arg Glu Gln Val Leu Leu Tyr Leu Phe Ala Leu
65 70 75 80

His Asp Tyr Asp Gln Ser Gly Gln Leu Asp Gly Leu Glu Leu Leu Ser
85 90 95

Met Leu Thr Ala Ala Leu Ala Pro Gly Ala Ala Asn Ser Pro Thr Thr
100 105 110

Asn Pro Val Ile Leu Ile Val
115

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 80 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA 60

AAAAAAAAAA AAAAAAAAAA 80

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2161 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

AGACAGGGAA TACTTTATTC AAAACCCATC ACAGAAATGG ACAGCTTGGG TCTGTAACAA 60

AGCATTCATG TTTTAGAGCA TAGGTCAGTA ATTGTATATG AGAGCATACA CTGCTACATA 120

CAAATTAACT GATCAGACCA CAACTTTTCA ATGTTTAAAA CAGAATAAGC TTCCCTGTAA 180

AAGCAGCACC TTTGTGACGT TTTAACTTTA GTATTCCTCT CCTTCTTCCT CACCCTCTCC 240

TTCAACAGAA TCCACACCAA CCTCCTCATA ATCCTTCCTC GCAGCACATG AATCACAGGT 300

ATTCCTACTG CAAGCGGGAG GCGGARGARC GGGAAGCGGC GGARCGCGAR GCGCGCGAGA 360

AAGGGCACTT GGAACCCACC GAGCTGCTGA TGAACCGGGC TTACTTGCAG AGCATTACCC 420

CTCAGGGGTA CTCTGACTCG GAGGAGAGGG AGAGTATGCC GAGGGATGGC GAGAGCGAGA 480

AGGAGCACGA GAAAGAAGGC GAGGATGGCT ACGGGAAGCT GGGCAGACAG GATGGCGACG 540

AGGAGTTCGA GGAGGAAGAG GAAGAAAGTG AAAATAAAAG TATGGATACG GATCCCGAAA 600

CGATACGAGA TGAAAAAGAG ACTGGAGATC ACTCCATGGA CGATAGTTCG GAGGATGGGA 660

AAATGGAAAC CAAATCAGAC CACGAGGAAG ACAATATGGA AGATGGCATG TAATAAACTA 720

CTGCATTTTA AGCTTCCTAT TTTTTTTTCC AGTAGTATTG TTACCTGCTT GAAAACACTG 780

CTGTGTTAAG CTGTTCATGC ACGTGCCTGA CGCTTCCAGG AAGCTGTAGA GAGGGACAGA 840

AGGGGCGGTT CAGCCAAGAC AGATGTWGAC GGAGTTGGAG CTGGGTATTG TTAAAAACTG 900

CATTATGCAA AAATTTTGTA CAGTGTTAAG GCCTAAAAAC TGTGTGGTTC AGAGACTAAT 960

TCCTGTGTTT AATAGCATTT ATACTTTAAG CACAACTAGA AAATTGTAAG AATTGCACTC 1020

TACTTATGTA TCACTACAAA CTTTAAAAAA CTATGTCTAA TTTATATTAA TACATTTTAA 1080

AAAGGTGCCC GCACTACCAT ACATCAGTAT TTTTATTATT ATTATTGTTA TTCCTTTTTA 1140

ATTTAATGTG CTCGCACTAC AATGCATCAG TATTATGATT CCTCTGTACT TTCCTTTCGC 1200

TATTCATCAA TTTCCCATTT TTTTTTTCAG CTTAAGTAAC CACACAATTT TAGGCCTCAA 1260

TTTTTTTTTT TCTGTGAAGG AACTTGAAGT GATGCATGTG TGAATTTAAG ATACCGAAGT 1320

CTTAAAGTGA CCTGGACGTG AAGGAAAAAG TAAGATGAGA AATAAAGAAA GCCTTTGTAA 1380

GGTGGTTTTA AAAGCCTTAT ATGCAAACCT TTTAATCTGT GTTTCTGCAA GTGCCATCCT 1440

TGTACAGTGT TAAGAGGGTA ACATGGGTTA CCTTTGCACC AGCTTCAGTG TTAAGCTCAC 1500

CCTGTTCTTT GAAGCACCCA TGTCAGTATT AGAAGAATAG GCAGCAGTTC CTTAGTTTAC 1560

ATATGTTTGT GCAATTATTT TCTGTACTTT TTTGTTCATT AATTTTGTCA GTATTACACC 1620

AAACTGTTTT TGCAACAAAA AAATTTTTTT TGCATTCATT TAATTTTAGG TCAAATAACA 1680

TTTTATTTAT GTGGCTCATT TTATATTTCC TAATTTTATT TATTTCATAC TGTAGTGTAC 1740

AGTATTATAG TTCTTCAATA TATAGATATA TTTTAGTAAA AAAGGAACAT GACGTTGATC 1800

ATTTGGGCAA ATTTTACGTA AAGAGAAGAG CATTTATTGT GTTTTGGAAC ATTAATTGTG 1860

AGATGGGATT TTTCAATTTT ATTATTTWAT TTTTGTTTTT TTCCAATTAC TGGAAATTCC 1920

AAATTTGGGA ACTTTTGATA CGATCTTGTG AAAACACTGT ATTTTCGACT GAAAATTCCA 1980

CTTTCTTCAT CTTGTTTTTT AGCTAAAAAG AGGGACTGTT AAATACAATG TATGATACCA 2040

TGACAAAAAT CTTTCCTGAA TTGTCCTTTG TAAAAGTATT ATTGAATTTT CAATTTGTAA 2100

TTTCTTTTGA AAATGACCAT GCTCGAATAA AAATGTAGCC AAACTAAAAA AAAAAAAAAA 2160

A 2161

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 141 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

```
Met Asn His Arg Tyr Ser Tyr Cys Lys Arg Glu Ala Glu Glu Arg Glu
1               5                  10                  15

Ala Ala Glu Arg Glu Ala Arg Glu Lys Gly His Leu Glu Pro Thr Glu
            20                  25                  30

Leu Leu Met Asn Arg Ala Tyr Leu Gln Ser Ile Thr Pro Gln Gly Tyr
        35                  40                  45

Ser Asp Ser Glu Glu Arg Glu Ser Met Pro Arg Asp Gly Glu Ser Glu
    50                  55                  60

Lys Glu His Glu Lys Glu Gly Glu Asp Gly Tyr Gly Lys Leu Gly Arg
65                  70                  75                  80

Gln Asp Gly Asp Glu Glu Phe Glu Glu Glu Glu Glu Glu Ser Glu Asn
                85                  90                  95

Lys Ser Met Asp Thr Asp Pro Glu Thr Ile Arg Asp Glu Lys Glu Thr
            100                 105                 110

Gly Asp His Ser Met Asp Asp Ser Ser Glu Asp Gly Lys Met Glu Thr
        115                 120                 125

Lys Ser Asp His Glu Glu Asp Asn Met Glu Asp Gly Met
    130                 135                 140
```

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 2169 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

```
GCAGTTACTG GGARGGGGCT TGCTGTGGCC CTGTCAGGAA RARTAGAGCT CTGGTCCAGC     60

TCCGCGCAGG GAGGGAGGCT GTCACCATGC CGGCCTGCTG CAGCTGCAGT GATGTTTTCC     120

AGTATGAGAC GAACAAAGTC ACTCGGATCC AGAGCATGAA TTATGGCACC ATTAAGTGGT     180

TCTTCCACGT GATCATCTTT TCCTACGTTT GCTTTGCTCT GGTGAGTGAC AAGCTGTACC     240

AGCGGAAAGA GCCTGTCATC AGTTCTGTGC ACACCAAGGT GAAGGGGATA GCAGARGTGA     300

AAGAGGAGAT CGTGGAGAAT GGAGTGAAGA AGTTGGTGCA CAGTGTCTTT GACACCGCAG     360
```

ACTACACCTT CCCTTTGCAG GGGAACTCTT TCTTCGTGAT GACAAACTTT CTCAAAACAG 420

AAGGCCAAGA GCAGCGGTTG TGTCCCGAGT ATCCCACCCG CAGGACGSTS TGTTCYTCTG 480

ACCGAGGTTG WAAAAAGGGA TGGATGGACC CGCAGAGCAA AGGAATTCAG ACCGGAAGGT 540

GTGTAGTGCA TGAAGGGAAC CAGAAGACYT GTGAAGTCTY TGCCTGGWGC CCCATSGAGG 600

CAGTGGAAGA GGCCCCCCGG CCTGCTYTCT TGAACAGTGC CGAAAACTTC ACTGTGCTCA 660

TCAAGAACAA TATCGACTTC CCCGGCCACA ACTACACCAC GAGAAACATC CTGCCAGGTT 720

TAAACATCAC TTGTACCTTC CACAAGACTC AGAATCCACA GTGTCCCATT TTCCGACTAG 780

GAGACATCTT CCGAGAAACA GGCGATAATT TTTCAGATGT GGCAATTCAG GGCGGAATAA 840

TGGGCATTGA GATCTACTGG GACTGCAACC TAGACCGTTG GTTCCATCAC TGCCATCCCA 900

AATACAGTTT CCGTCGCCTT GACGACAAGA CCACCAACGT GTCCTTGTAC CCTGGCTACA 960

ACTTCAGATA CGCCAAGTAC TACAAGGAAA ACAATGTTGA GAAACGGACT CTGATAAAAG 1020

TCTTCGGGAT CCGTTTTGAC ATCCTGGTTT TTGGCACCGG AGGAAAATTT GACATTATCC 1080

AGCTGGTTGT GTACATCGGC TCAACCCTCT CCTACTTCGG TCTGGCCGCT GTGTTCATCG 1140

ACTTCCTCAT CGACACTTAC TCCAGTAACT GCTGTCGCTC CCATATTTAT CCCTGGTGCA 1200

AGTGCTGTCA GCCCTGTGTG GTCAACGAAT ACTACTACAG GAAGAAGTGC GAGTCCATTG 1260

TGGAGCCAAA GCCGACATTA AAGTATGTGT CCTTTGTGGA TGAATCCCAC ATTAGGATGG 1320

TGAACCAGCA GCTACTAGGG AGAAGTCTGC AAGATGTCAA GGGCCAAGAA GTCCCAAGAC 1380

CTGCGATGGA CTTCACAGAT TTGTCCAGGC TGCCCCTGGC CCTCCATGAC ACACCCCCGA 1440

TTCCTGGACA ACCAGAGGAG ATACAGCTGC TTAGAAAGGA GGCGACTCCT AGATCCAGGG 1500

ATAGCCCCGT CTGGTGCCAG TGTGGAAGCT GCCTCCCATC TCAACTCCCT GAGAGCCACA 1560

GGTGCCTGGA GGAGCTGTGC TGCCGGAAAA AGCCGGGGGC CTGCATCACC ACCTCAGAGC 1620

TGTTCAGGAA GCTGGTCCTG TCCAGACACG TCCTGCAGTT CCTCCTGCTC TACCAGGAGC 1680

CCTTGCTGGC GCTGGATGTG GATTCCACCA ACAGCCGGCT GCGGCACTGT GCCTACAGGT 1740

GCTACGCCAC CTGGCGCTTC GGCTCCCAGG ACATGGCTGA CTTTGCCATC CTGCCCAGCT 1800

GCTGCCGCTG GAGGATCCGG AAAGAGTTTC CGAAGAGTGA AGGGCAGTAC AGTGGCTTCA 1860

AGAGTCCTTA CTGAAGCCAG GCACCGTGGC TMACGTCTGT AATCCCAGCG CTTTGGGAGG 1920

CCGAGGCAGG CAGATCACCT GAGGTCGGGA GTTGGAGACC CGCCTGGCTA ACAAGGCGAA 1980

ATCCTGTCTG TACTAAAAAT ACAAAAATCA GCCAGACATG GTGGCATGCA CCTGCAATCC 2040

CAGCTACTCG GGAGGCTGAG GCACAAGAAT CACTTGAACC CGGGAGGCAG AGGTTGTAGT 2100

GAGCCCAGAT TGTGCCACTG CTYTCCAGCC TGGGAGGCAC AGCAAACTGT CCCCAAAAAA 2160

AAAAAAAAA 2169

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 595 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

```
Met Pro Ala Cys Cys Ser Cys Ser Asp Val Phe Gln Tyr Glu Thr Asn
1               5                  10                  15

Lys Val Thr Arg Ile Gln Ser Met Asn Tyr Gly Thr Ile Lys Trp Phe
            20                  25                  30

Phe His Val Ile Ile Phe Ser Tyr Val Cys Phe Ala Leu Val Ser Asp
        35                  40                  45

Lys Leu Tyr Gln Arg Lys Glu Pro Val Ile Ser Ser Val His Thr Lys
    50                  55                  60

Val Lys Gly Ile Ala Glu Val Lys Glu Glu Ile Val Glu Asn Gly Val
65                  70                  75                  80

Lys Lys Leu Val His Ser Val Phe Asp Thr Ala Asp Tyr Thr Phe Pro
                85                  90                  95

Leu Gln Gly Asn Ser Phe Phe Val Met Thr Asn Phe Leu Lys Thr Glu
            100                 105                 110

Gly Gln Glu Gln Arg Leu Cys Pro Glu Tyr Pro Thr Arg Arg Thr Xaa
        115                 120                 125

Cys Ser Ser Asp Arg Gly Xaa Lys Lys Gly Trp Met Asp Pro Gln Ser
    130                 135                 140

Lys Gly Ile Gln Thr Gly Arg Cys Val Val His Glu Gly Asn Gln Lys
145                 150                 155                 160

Thr Cys Glu Val Xaa Ala Trp Xaa Pro Xaa Glu Ala Val Glu Glu Ala
                165                 170                 175

Pro Arg Pro Ala Xaa Leu Asn Ser Ala Glu Asn Phe Thr Val Leu Ile
            180                 185                 190
```

Lys Asn Asn Ile Asp Phe Pro Gly His Asn Tyr Thr Thr Arg Asn Ile
195 200 205

Leu Pro Gly Leu Asn Ile Thr Cys Thr Phe His Lys Thr Gln Asn Pro
210 215 220

Gln Cys Pro Ile Phe Arg Leu Gly Asp Ile Phe Arg Glu Thr Gly Asp
225 230 235 240

Asn Phe Ser Asp Val Ala Ile Gln Gly Gly Ile Met Gly Ile Glu Ile
245 250 255

Tyr Trp Asp Cys Asn Leu Asp Arg Trp Phe His His Cys His Pro Lys
260 265 270

Tyr Ser Phe Arg Arg Leu Asp Asp Lys Thr Thr Asn Val Ser Leu Tyr
275 280 285

Pro Gly Tyr Asn Phe Arg Tyr Ala Lys Tyr Tyr Lys Glu Asn Asn Val
290 295 300

Glu Lys Arg Thr Leu Ile Lys Val Phe Gly Ile Arg Phe Asp Ile Leu
305 310 315 320

Val Phe Gly Thr Gly Gly Lys Phe Asp Ile Ile Gln Leu Val Val Tyr
325 330 335

Ile Gly Ser Thr Leu Ser Tyr Phe Gly Leu Ala Ala Val Phe Ile Asp
340 345 350

Phe Leu Ile Asp Thr Tyr Ser Ser Asn Cys Cys Arg Ser His Ile Tyr
355 360 365

Pro Trp Cys Lys Cys Cys Gln Pro Cys Val Val Asn Glu Tyr Tyr Tyr
370 375 380

Arg Lys Lys Cys Glu Ser Ile Val Glu Pro Lys Pro Thr Leu Lys Tyr
385 390 395 400

Val Ser Phe Val Asp Glu Ser His Ile Arg Met Val Asn Gln Gln Leu
405 410 415

Leu Gly Arg Ser Leu Gln Asp Val Lys Gly Gln Glu Val Pro Arg Pro
420 425 430

Ala Met Asp Phe Thr Asp Leu Ser Arg Leu Pro Leu Ala Leu His Asp
435 440 445

Thr Pro Pro Ile Pro Gly Gln Pro Glu Glu Ile Gln Leu Leu Arg Lys
450 455 460

Glu Ala Thr Pro Arg Ser Arg Asp Ser Pro Val Trp Cys Gln Cys Gly
465 470 475 480

Ser Cys Leu Pro Ser Gln Leu Pro Glu Ser His Arg Cys Leu Glu Glu
485 490 495

```
Leu Cys Cys Arg Lys Lys Pro Gly Ala Cys Ile Thr Thr Ser Glu Leu
            500                 505                 510

Phe Arg Lys Leu Val Leu Ser Arg His Val Leu Gln Phe Leu Leu Leu
        515                 520                 525

Tyr Gln Glu Pro Leu Leu Ala Leu Asp Val Asp Ser Thr Asn Ser Arg
    530                 535                 540

Leu Arg His Cys Ala Tyr Arg Cys Tyr Ala Thr Trp Arg Phe Gly Ser
545                 550                 555                 560

Gln Asp Met Ala Asp Phe Ala Ile Leu Pro Ser Cys Cys Arg Trp Arg
                565                 570                 575

Ile Arg Lys Glu Phe Pro Lys Ser Glu Gly Gln Tyr Ser Gly Phe Lys
            580                 585                 590

Ser Pro Tyr
        595
```

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

GNGTGAAGTCT TCCTGGGAAC CATAATCT 29

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

TNTTCCCTGAA GAGCTGGAGA GGTGCTAA 29

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

GNTCTACCATG TGAAGAAGGA ACGCAAAA 29

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

TNGGCAAAGCT GTGCTGCAGG ATAGAGTG 29

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

ANGTGAAGGAC TGTAAGCCCA CACAAGCT 29

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

TNATGACAGCC CAGGCTGTGT TATCTGCA 29

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

CNGGGCCTTCG TTTTCACTAC CTTCAGAA 29

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

TNAAGACGGTG ACTGAGAGGC AGACAAGG 29

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs

(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

TNCCTAGTCCC TTTAGGTAGC TCTGCAGA 29

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

GNGCACGAGAA AGAAGGCGAG GATGGCTA 29

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

TNGGACACTGT GGATTCTGAG TCTTGTGG 29

What is claimed is:

1. A composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 437 to nucleotide 1159;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 515 to nucleotide 1159;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 539 to nucleotide 1099;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone AR415_4 deposited under accession number ATCC 98232;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone AR415_4 deposited under accession number ATCC 98232;

(g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone AR415_4 deposited under accession number ATCC 98232;

(h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone AR415_4 deposited under accession number ATCC 98232;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:2;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:2 having biological activity;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

2. A composition of claim 1 wherein said polynucleotide is operably linked to an expression control sequence.

3. A host cell transformed with a composition of claim 2.

4. The host cell of claim 3, wherein said cell is a mammalian cell.

5. A process for producing a protein, which comprises:

(a) growing a culture of the host cell of claim 3 in a suitable culture medium; and

(b) purifying the protein from the culture.

6. A protein produced according to the process of claim 5.

7. The protein of claim 6 comprising a mature protein.

8. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:2;

(b) the amino acid sequence of SEQ ID NO:2 from amino acid 51 to amino acid 221;

(c) fragments of the amino acid sequence of SEQ ID NO:2; and

(d) the amino acid sequence encoded by the cDNA insert of clone AR415_4 deposited under accession number ATCC 98232;

the protein being substantially free from other mammalian proteins.

9. The composition of claim 8, wherein said protein comprises the amino acid sequence of SEQ ID NO:2.

10. The composition of claim 8, wherein said protein comprises the amino acid sequence of SEQ ID NO:2 from amino acid 51 to amino acid 221.

11. The composition of claim 8, further comprising a pharmaceutically acceptable carrier.

12. A method for preventing, treating or ameliorating a medical condition which comprises administering to a mammalian subject a therapeutically effective amount of a composition of claim 11.

13. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:1.

14. A composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 59 to nucleotide 376;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 179 to nucleotide 376;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone AS63_29 deposited under accession number ATCC 98232;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone AS63_29 deposited under accession number ATCC 98232;

(f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone AS63_29 deposited under accession number ATCC 98232;

(g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone AS63_29 deposited under accession number ATCC 98232;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:4;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:4 having biological activity;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

15. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:4;

(b) the amino acid sequence of SEQ ID NO:4 from amino acid 1 to amino acid 91;

(c) fragments of the amino acid sequence of SEQ ID NO:4; and

(d) the amino acid sequence encoded by the cDNA insert of clone AS63_29 deposited under accession number ATCC 98232;

the protein being substantially free from other mammalian proteins.

16. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:3 or SEQ ID NO:5.

17. A composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:6;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:6 from nucleotide 198 to nucleotide 2039;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:6 from nucleotide 490 to nucleotide 809;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone AY304_14 deposited under accession number ATCC xxxxx;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone AY304_14 deposited under accession number ATCC xxxxx;

(f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone AY304_14 deposited under accession number ATCC xxxxx;

(g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone AY304_14 deposited under accession number ATCC xxxxx;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:7;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:7 having biological activity;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

18. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:7;

(b) the amino acid sequence of SEQ ID NO:7 from amino acid 126 to amino acid 204;

(c) the amino acid sequence of SEQ ID NO:7 from amino acid 106 to amino acid 204;

(d) fragments of the amino acid sequence of SEQ ID NO:7; and

(e) the amino acid sequence encoded by the cDNA insert of clone AY304_14 deposited under accession number ATCC xxxxx;

the protein being substantially free from other mammalian proteins.

19. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:6.

20. A composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:8;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:8 from nucleotide 102 to nucleotide 2027;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:8 from nucleotide 1902 to nucleotide 2027;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:8 from nucleotide 1 to nucleotide 431;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone BG160_1 deposited under accession number ATCC 98232;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone BG160_1 deposited under accession number ATCC 98232;

(g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone BG160_1 deposited under accession number ATCC 98232;

(h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone BG160_1 deposited under accession number ATCC 98232;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:9;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:9 having biological activity;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

21. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:9;

(b) the amino acid sequence of SEQ ID NO:9 from amino acid 1 to amino acid 110;

(c) fragments of the amino acid sequence of SEQ ID NO:9; and

(d) the amino acid sequence encoded by the cDNA insert of clone BG160_1 deposited under accession number ATCC 98232;

the protein being substantially free from other mammalian proteins.

22. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:8.

23. A composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:11;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:11 from nucleotide 566 to nucleotide 631;

(c) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone BO432_4 deposited under accession number ATCC 98232;

(d) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone BO432_4 deposited under accession number ATCC 98232;

(e) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone BO432_4 deposited under accession number ATCC 98232;

(f) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone BO432_4 deposited under accession number ATCC 98232;

(g) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:12;

(h) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:12 having biological activity;

(i) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(f) above;

(j) a polynucleotide which encodes a species homologue of the protein of (g) or (h) above ; and

(k) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(h).

24. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:12;

(b) fragments of the amino acid sequence of SEQ ID NO:12; and

(c) the amino acid sequence encoded by the cDNA insert of clone BO432_4 deposited under accession number ATCC 98232;

the protein being substantially free from other mammalian proteins.

25. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:11, SEQ ID NO:10 or SEQ ID NO:13 .

26. A composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:14;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:14 from nucleotide 45 to nucleotide 428;

(c) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone BO538_2 deposited under accession number ATCC 98232;

(d) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone BO538_2 deposited under accession number ATCC 98232;

(e) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone BO538_2 deposited under accession number ATCC 98232;

(f) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone BO538_2 deposited under accession number ATCC 98232;

(g) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:15;

(h) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:15 having biological activity;

(i) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(f) above;

(j) a polynucleotide which encodes a species homologue of the protein of (g) or (h) above ; and

(k) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(h).

27. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:15;

(b) the amino acid sequence of SEQ ID NO:15 from amino acid 52 to amino acid 128;

(c) fragments of the amino acid sequence of SEQ ID NO:15; and

(d) the amino acid sequence encoded by the cDNA insert of clone BO538_2 deposited under accession number ATCC 98232;

the protein being substantially free from other mammalian proteins.

28. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:14 or SEQ ID NO:16.

29. A composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17 from nucleotide 144 to nucleotide 566;

(c) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone BR595_4 deposited under accession number ATCC 98232;

(d) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone BR595_4 deposited under accession number ATCC 98232;

(e) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone BR595_4 deposited under accession number ATCC 98232;

(f) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone BR595_4 deposited under accession number ATCC 98232;

(g) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:18;

(h) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:18 having biological activity;

(i) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(f) above;

(j) a polynucleotide which encodes a species homologue of the protein of (g) or (h) above ; and

(k) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(h).

30. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:18;

(b) the amino acid sequence of SEQ ID NO:18 from amino acid 39 to amino acid 141;

(c) fragments of the amino acid sequence of SEQ ID NO:18; and

(d) the amino acid sequence encoded by the cDNA insert of clone BR595_4 deposited under accession number ATCC 98232;

the protein being substantially free from other mammalian proteins.

31. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:17 or SEQ ID NO:19.

32. A composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:20;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:20 from nucleotide 232 to nucleotide 1041;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:20 from nucleotide 460 to nucleotide 1041;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:20 from nucleotide 590 to nucleotide 1163;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CI490_2 deposited under accession number ATCC 98232;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CI490_2 deposited under accession number ATCC 98232;

(g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CI490_2 deposited under accession number ATCC 98232;

(h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CI490_2 deposited under accession number ATCC 98232;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:21;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:21 having biological activity;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

33. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:21;

(b) the amino acid sequence of SEQ ID NO:21 from amino acid 133 to amino acid 270;

(c) fragments of the amino acid sequence of SEQ ID NO:21; and

(d) the amino acid sequence encoded by the cDNA insert of clone CI490_2 deposited under accession number ATCC 98232;

the protein being substantially free from other mammalian proteins.

34. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:20.

35. A composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:22;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:22 from nucleotide 268 to nucleotide 624;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:22 from nucleotide 325 to nucleotide 624;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CI522_1 deposited under accession number ATCC 98232;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CI522_1 deposited under accession number ATCC 98232;

(f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CI522_1 deposited under accession number ATCC 98232;

(g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CI522_1 deposited under accession number ATCC 98232;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:23;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:23 having biological activity;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

36. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:23;

(b) fragments of the amino acid sequence of SEQ ID NO:23; and

(c) the amino acid sequence encoded by the cDNA insert of clone CI522_1 deposited under accession number ATCC 98232;

the protein being substantially free from other mammalian proteins.

37. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:22 or SEQ ID NO:24.

38. A composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:25;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:25 from nucleotide 288 to nucleotide 713;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:25 from nucleotide 686 to nucleotide 968;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CN238_1 deposited under accession number ATCC 98232;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CN238_1 deposited under accession number ATCC 98232;

(f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CN238_1 deposited under accession number ATCC 98232;

(g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CN238_1 deposited under accession number ATCC 98232;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:26;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:26 having biological activity;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

39. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:26;

(b) fragments of the amino acid sequence of SEQ ID NO:26; and

(c) the amino acid sequence encoded by the cDNA insert of clone CN238_1 deposited under accession number ATCC 98232;

the protein being substantially free from other mammalian proteins.

40. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:25.

41. A composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:27;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:27 from nucleotide 87 to nucleotide 1874;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:27 from nucleotide 452 to nucleotide 830;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CO390_1 deposited under accession number ATCC 98232;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CO390_1 deposited under accession number ATCC 98232;

(f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CO390_1 deposited under accession number ATCC 98232;

(g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CO390_1 deposited under accession number ATCC 98232;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:28;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:28 having biological activity;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

42. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:28;

(b) the amino acid sequence of SEQ ID NO:28 from amino acid 140 to amino acid 248;

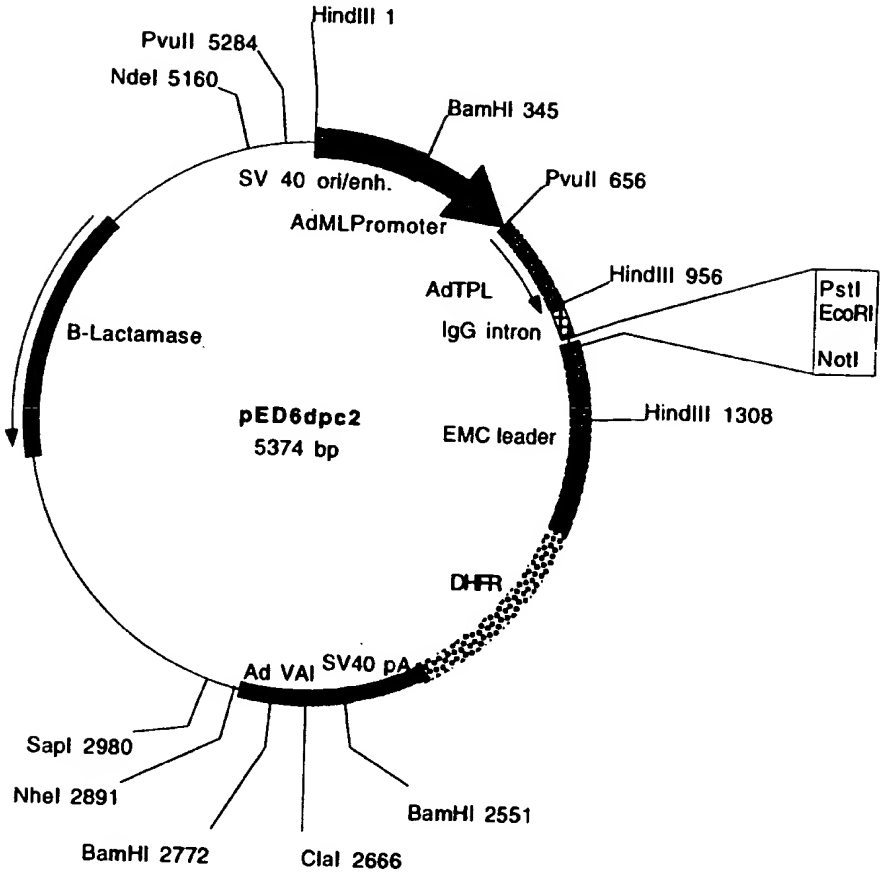
(c) fragments of the amino acid sequence of SEQ ID NO:28; and

(d) the amino acid sequence encoded by the cDNA insert of clone CO390_1 deposited under accession number ATCC 98232;

the protein being substantially free from other mammalian proteins.

43. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:27.

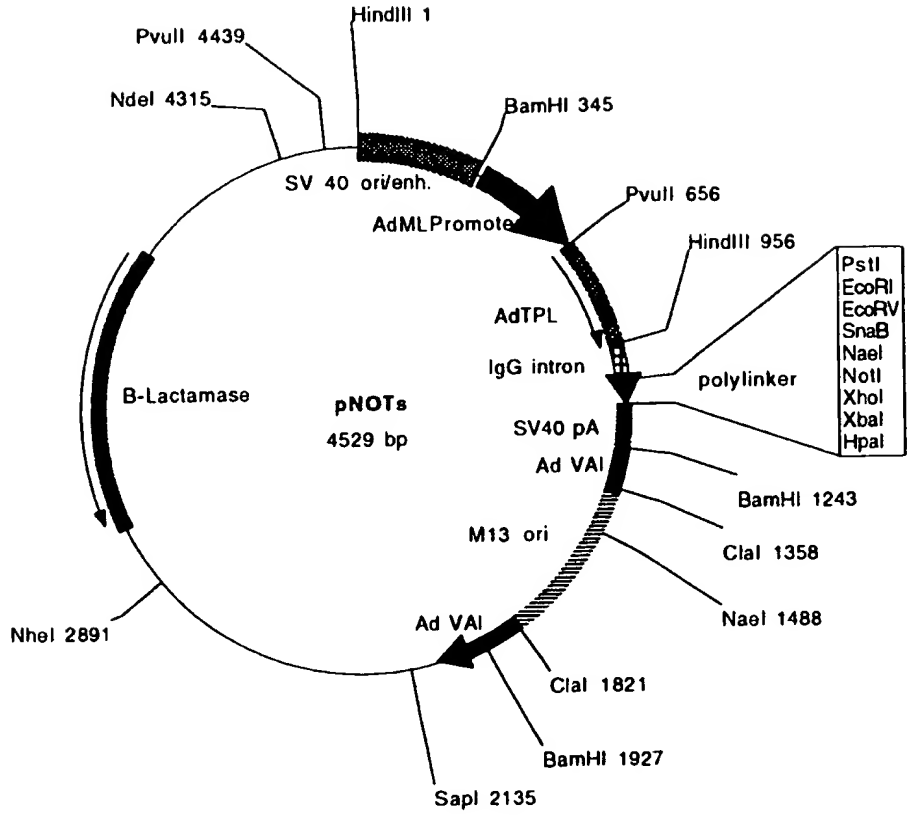
# FIGURE 1A




**Plasmid name:** pED6dpc2
**Plasmid size:** 5374 bp

**Comments/References:** pED6dpc2 is derived from pED6dpc1 by insertion of a new polylinker to facilitate cDNA cloning. SST cDNAs are cloned between EcoRI and NotI. pED vectors are described in Kaufman et al.(1991), NAR 19: 4485-4490.

# FIGURE 1B




**Plasmid name:** pNOTs
**Plasmid size:** 4529 bp

**Comments/References:** pNOTs is a derivative of pMT2 (Kaufman et al,1989. Mol.Cell.Biol.9:1741-1750). DHFR was deleted and a new polylinker was inserted between EcoRI and HpaI. M13 origin of replication was inserted in the ClaI site. SST cDNAs are cloned between EcoRI and NotI